

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
1 July 2004 (01.07.2004)

PCT

(10) International Publication Number
WO 2004/055216 A2

(51) International Patent Classification⁷:

C12Q 1/70

(21) International Application Number:

PCT/US2003/039722

(22) International Filing Date:

12 December 2003 (12.12.2003)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

60/433,303 13 December 2002 (13.12.2002) US

(71) Applicant (for all designated States except US): FOX CHASE CANCER CENTER [US/US]; 333 Cottman Avenue, Philadelphia, PA 19111 (US).

(74) Agents: RIGAET, Kathleen, D. et al.; Dann, Dorfman, Herrell & Skillman, Suite 2400, 1601 Market Street, Philadelphia, PA 19103-2307 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

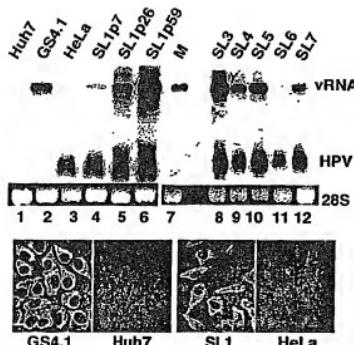
(84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ,UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: REPLICATION OF HEPATITIS C VIRUS IN NON-HEPATIC EPITHELIAL AND MOUSE HEPATIC CELLS



(57) Abstract: Cells and cell lines which replicate HCV of non-hepatic human and non-human origin are disclosed. Also provided are methods of using such cells and cell lines to identify anti-HCV agents for the treatment of HCV infection.

**REPLICATION OF HEPATITIS C VIRUS IN NON-HEPATIC
EPITHELIAL AND MOUSE HEPATIC CELLS**

Qing Zhu

Ju-Tao Guo

5

Christoph Seeger

This application claims priority to U.S. provisional application 60/433,303, filed December 13, 2002, the entire contents of which are incorporated by reference
10 herein.

GOVERNMENT RIGHT

Pursuant to 35 U.S.C. Section 202(c), it is acknowledged that the United States government has
15 certain rights in the invention described herein, which was made in part with funds from the National Institutes of Health Grant No. AI48046.

FIELD OF THE INVENTION

20 This invention relates to the fields of molecular biology and pathology. Novel animal cell lines and non-hepatic human epithelial cell lines for the replication of hepatitis C virus (HCV), as well as methods for screening for anti-HCV drugs or HCV receptors using these
25 cell lines are disclosed. Furthermore, adaptive sequence mutations in the HCV genome, which permit replication in non-human, and non-hepatic cell lines are also provided.

BACKGROUND OF THE INVENTION

30 Several publications and patent documents are cited in this application in order to more fully describe the state of the art to which this invention pertains. The disclosure of each of these citations is incorporated by reference herein.

35 Hepatitis C virus (HCV) is an enveloped, positive

stranded RNA virus that belongs to the *Flaviviridae*, a family of viruses including human pathogens such as yellow fever virus, dengue virus and West Nile virus (Q. L. Choo *et al.*, *Science* 244, 359-62 (1989)). Although broad tissue and species tropisms are hallmarks of these viruses, HCV replication has so far only been detected in human and chimpanzee livers. Moreover, for reasons that are not yet understood, HCV RNA levels in infected liver tissue are extremely low, generally below one copy of RNA per cell and hence, can only be detected with PCR, making it difficult to determine whether secondary sites for viral replication exist (J. Boisvert *et al.*, *J Infect Dis* 184, 827-35 (Oct 1, 2001); R. E. Lanford, *et al.*, *J Virol* 69, 8079-83 (1995)).

HCV encodes a polyprotein that is processed proteolytically into ten polypeptides (K. E. Reed, C. M. Rice, *Curr Top Microbiol Immunol* 242, 55-84 (2000)). Three of them are structural proteins required for capsid formation (core) and assembly into enveloped viral particles (E1 and E2). Four of them are enzymes including cysteine and serine proteases (NS2 and NS3), an ATP dependent helicase (NS3) and a RNA-directed RNA polymerase (NS5B). The functions of the remaining three polypeptides, p7, NS4B, and NS5A, for viral replication are not yet known. For study of replication of HCV in tissue culture cells, the structural proteins can be replaced with a selectable marker, such as the neomycin phosphotransferase. See for example Figure 2, left panel of Lohman *et al.* (V. Lohmann *et al.*, *Science* 285, 110-3 (1999)). Replication of such subgenomic HCV replicons in tissue culture cells has so far only been demonstrated in the human hepatoma cell line Huh7, consistent with the narrow host and tissue tropism of HCV infections.

HCV infection poses a significant public health problem. Approximately 3% of the world's population has

5 persistent HCV infection. In 1989, the virus was identified as the major aetiological agent responsible for post-transfusion non-A and non-B hepatitis.

10 Following primary HCV infection, persistent viraemia and chronic hepatitis develop in the majority of cases. Efforts to elucidate the mechanisms behind viral persistence and hepatocellular damage have been frustrated by the lack of a reliable cell culture system for viral propagation *in vitro*. In addition, as the 15 chimpanzee is the only experimental animal susceptible to HCV infection, progress in research is hampered by the lack of a small animal model to facilitate pathophysiological studies as well as the evaluation of antiviral treatment and vaccine strategies.

20 15 Furthermore, although the initial HCV infection is asymptomatic, subsequent clinical manifestations of HCV induced liver disease include fibrosis, cirrhosis, and hepatocellular carcinoma (Alter, H. J., and L. B. Seeff. 2000. *Semin. Liver Dis.* 20:17-35). Combination antiviral 25 therapy with alpha interferon (IFN- α) and ribavirin, a purine nucleoside analogue, arrests disease progression and can lead to sustained recovery in only 45 to 80% of treated patients (Di Bisceglie, A. M., and J. H. Hoofnagle. 2002. *Hepatology* 36:S121-S127). Additionally, response to IFN- α therapy can vary significantly depending on the viral genotype, ranging from 30 to 40% for genotype 1 to as high as 80% for genotypes 2 and 3. This suggests that viral determinants also play an important role in regulating the cellular IFN response 30 against HCV (Kinzie, J. L., et al., 2001. *J. Viral Hepatitis* 8:264-269; McHutchison, J. G., et al., 1998. *N. Engl. J. Med.* 339:1485-1492). The parameters determining the success or failure of antiviral therapy are not understood, and their identification represents a major 35 challenge in HCV biology.

Therefore, there is a desperate need for non-hepatic cell culture systems, and small animal models for the identification and characterization of anti-viral agents for the prevention and treatment of HCV infection.

- 5 Additionally, there is a need in the art to elucidate the mechanism of HCV inhibition by IFN- α , so that other treatments may be found.

SUMMARY OF THE INVENTION

10 The present invention provides HCV replicating cells and cell lines derived from human non-hepatic cells or non-human cells. According to one embodiment of the invention, the cells are human epithelial cells of non-liver origin, such as, HeLa cells. According to another 15 embodiment of the invention, the cells capable of replicating HCV are hepatoma and hepatocyte cells of mouse origin, such as, Hepal-6 cells, or AML12 cells respectively.

The present invention also provides a non-human host 20 animal comprising cells infected with HCV. In one embodiment of the invention, the host animal is a mouse. In another embodiment of the invention, the cells infected with HCV are mouse hepatoma cells.

Also provided by the present invention are methods 25 for producing human non-hepatic cells or non-human cells that are capable of replicating HCV, and cell lines comprising the same. Such methods include transfection with total HCV RNA or an HCV replicon which comprises one or more adaptive mutations which facilitate replication 30 in a cell of interest.

The present invention further provides methods for screening an agent that modulates HCV replication by incubating the agent with the aforementioned cells or administering the agent to the aforementioned host animal 35 comprising cells replicating HCV and assessing said agent

for modulation of HCV replication. Such agents may inhibit or enhance production of HCV. These agents may be cytopathic or non-cytopathic to HCV infected cells. Agents which activate aspects of the JAK/STAT pathway may 5 also be screened using the cells and cell lines of the invention.

Also provided by the present invention are HCV derived polynucleotides comprising adaptive mutations. The present inventor has discovered that these mutations 10 are associated with expanded tropism of HCV.

Additionally, the present invention provides polypeptides encoded by the mutated HCV polynucleotides described above.

15 BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are a Northern blot (Fig. 1A) and micrographs showing replication of HCV subgenomic replicons in HeLa cells. (Fig. 1A) Detection of HCV viral RNA. Total RNA (5 µg) was isolated from HeLa cell 20 lines (SL1, SL3-7) that were established from G418 resistant cell colonies and analyzed by Northern blot analysis. Blots were hybridized with radiolabeled RNA probes corresponding to the HCV NS5 region to detect viral RNA (vRNA) and the ΔE1 region of human papilloma 25 virus (HPV) present in HeLa cells. In vitro transcribed HCV RNA (1 ng, + 5µg total RNA from Huh7 cells, lane 7) served as a marker (M) and control for the hybridization reaction and 28S ribosomal RNA as a control for the amount of RNA present in each sample analyzed. GS4.1 is a 30 Huh7 derived cell line expressing HCV subgenomic replicons. RNA in SL1 cells was analyzed from cells harvested at the indicated passage (p). SL3-7 were analyzed at p3 (Fig. 1B). Immunohistochemical analysis of HCV replication in HeLa cells. Expression of NS5A in 35 GS4.1 and SL1 cells (p26) was detected with a monoclonal

antibody bound to fluorescent isothiocyanate (FITC)-conjugated antibody. Parental Huh7 and HeLa cells served as controls.

Figure 2 diagrams sequence analysis of HCV subgenomes in HeLa and mouse hepatoma Hepal-6 cells. (Left section) The physical map of HCV subgenomic RNA include the positions of the first amino acid NS3 and the last residue of the polyprotein. The internal ribosomal entry site for the translation of the NS genes is indicated (EMCV-IRES). (Right section) Mutations causing amino acid changes identified in cDNAs isolated from subgenomic replicons present in GS4.1 and the indicated HeLa (SL1 (p26); SL2, p5) and Hepal-6 cell lines (MH1 (p12), MH2 (p4), MH4 (p4)) are depicted with horizontal bars. Four independent clones were sequenced from each PCR fragment that was amplified from cDNAs obtained from total RNA purified from the indicated cell lines. Mutations present in more than one cell line are indicated by amino acid position. Mutations that occurred in 50% of the clones analyzed are marked with an asterisk. Mutations that occurred in only one of four clones analyzed were not included in the figure. A deletion identified in cDNA clones obtained from SL1 cells spanning amino acids 2371 to 2413 is indicated (Δ).

Figures 3A and 3B are a Northern blot and micrographs showing replication of HCV subgenomes in mouse hepatoma cells. (A) Detection of HCV RNA. RNA in Hepal-6 cell lines (MH1-5) that were established from G418 resistant cell colonies and analyzed as described in the legend to Figure 1 except that a radiolabeled probe specific to mouse albumin cDNA was used in lieu of the probe against HPV. MH4-5 were analyzed at p4 (B). Immunohistochemical analysis of HCV replication in MH1 cells (p3). Expression of NS5A was detected as described in Figure 1. Hepal-6 cells served as a

negative control.

Figure 4 is a Northern blot showing detection of HCV RNA in mouse hepatocyte cells. AML12 cells were transfected with total RNA from GS4.1 cells, which express subgenomic HCV replicons, and HFL cells, which express full length HCV genomes. 5 micrograms of total RNA was isolated from AML12 cell lines (MA6-5 to MA6-8, and MAC1) that were established from G418 resistant cell colonies and analyzed by Northern blot analysis. sgRNA indicates subgenomic RNA and f1RNA full length genomic RNA.

Figures 5A and 5B are two Northern blots showing the antiviral activity of the HCV RNA polymerase inhibitor 2'-C-methyladenosine (2CMA). GS4.1 (Huh7) cells (5A) and SL1 (HeLa) cells (5B) were treated with 10 μ M 2CMA. The cells were harvested at the indicated time points. Total cellular RNA was extracted and viral RNA (vRNA) analyzed by Northern blot analysis.

Figures 6A and 6B are two graphs depicting antiviral activity of the HCV RNA polymerase inhibitor 5-OH-cytidine. GS4.1 (Huh7) cells and SL1 (HeLa) cells were treated with the indicated amounts of 5-OH-cytidine. The DNA polymerase inhibitor 5-OH-deoxy-cytidine was used as a negative control. The cells were harvested 72 hours after incubation with the drugs. Total cellular RNA was extracted and viral RNA analyzed by Northern blot analysis. The intensity of the bands corresponding to HCV RNA was determined with a Fuji phosphoimager.

Figures 7A and 7B are a Northern blot and a graph showing antiviral activity of IFN- α in Huh 7 (GS4.1) and HeLa (SL1) cell lines containing HCV replicons. (7A) Viral RNA (vRNA) levels present in GS4.1 and SL1 cells incubated with 0, 0.1, 0.3, 1, 3, 10, 30, and 100 IU of IFN- α /ml (lanes 1 to 9 and 12 to 18) and with 0.01 and 0.03 IU/ml (lanes 10 and 11) for 72 h were determined by

Northern blot analysis. rRNA (28S rRNA) served as a control for the amount of RNA loaded per lane. (7B) Amounts of HCV RNA were determined with a Fuji phosphorimager, and the values were plotted as the 5 percentages of the values obtained with untreated cells in lanes 1 and 9.

Figures 8A and 8B are micrographs and histograms showing that IFN- α induces apoptosis of SL1 cells. (8A) Annexin V-FITC staining. SL1 cells grown on glass 10 coverslips were left untreated (upper left) or treated with 100 IU of IFN- α /ml for 6 h (upper right) or 20 h (lower left) or with 100 IU of IFN- α /ml and 20 μ M caspase inhibitor ZVAD-FMK for 20 h (lower right). Cells were then processed for annexin V FITC staining and 15 viewed with a fluorescence microscope. (8B) Flow cytometry analysis. SL1 cells were left untreated (upper left) or treated with 100 IU of IFN- α /ml (upper right) or with IFN- α and 20 μ M ZVAD-FMK (lower left) for 24 h. To inhibit viral replication, SL1 cells were incubated at 20 39°C for 60 h and then treated with 100 IU of IFN- α /ml for 24 h (lower left). Cells were harvested and processed for annexin V-FITC staining and analyzed by flow cytometry. The percentages of FITC-positive cells are indicated.

Figures 9A and 9B are Northern blots and a graph 25 showing a comparison of IFN- α responses against HCV and flavivirus Kunjin virus replicons in HeLa cells. (9A) SL1 and KUNCD20 cells were incubated with 0, 0.01, 0.04, 0.16, 0.625, 2.5, 10, 40, and 160 IU of IFN- α /ml (lanes 30 1 to 9 and 10 to 18, respectively) for 72 h, and viral RNA levels were determined by Northern blot analysis with a plus-strand-specific RNA probe for the neomycin phosphotransferase II gene. Mx-1 mRNA served as a control for IFN- α -induced gene expression. β -Actin mRNA served 35 as a control for the amount of RNA loaded per lane. (9B)

The amounts of HCV and Kunjin virus replicon RNA (arbitrary units) were determined with a Fuji phosphorimager.

Figure 10 shows an overview of the interactions between a virus and the IFN system. Replication of viruses in cells produces dsRNA and viral proteins, which activate PKR and OAS/RNase L antiviral pathways and also signal to the promoter of the IFN- β gene by activating transcription factors IRF3, NF- κ B, and ATF2. Secreted IFN binds to its receptor and activates receptor-associated Jak kinases, leading to the formation of the trimeric transcription factor ISGF3, which binds to the IFN-stimulated response element (ISRE) on promoters of IFN-stimulated genes. Among the products of the several hundred genes induced by IFN, PKR, OAS/RNase L, and Mx are the best-characterized antiviral proteins, which inhibit different stages of viral replication and induce apoptosis of virally infected cells.

Figures 11A-11D are Northern blots and two graphs showing inhibition of the IFN- α response by genistein and the V protein of HPIV2. (11A) GS4.1 cells were incubated with 100 μ g of genistein/ml for 2 h and then with 100 IU/ml IFN- α for an additional 24 h. Viral RNA levels were determined by Northern blot analysis. The cells were harvested at the indicated time points, and MxA mRNA and viral RNA levels were determined by Northern blot analysis. Ribosomal 28S RNA was used as a control for the amount of RNA loaded on each lane. (11B) The amounts of viral RNA were measured with a phosphorimager and plotted as percentages of the values obtained with untreated cells. (11C) GS4.1 cells were transfected with pCMV-E3L and pEF-HA-HPIV2 and treated with IFN- α at the indicated concentrations for 3 days. HCV RNA was subjected to Northern blot analysis. (11D) Viral RNA levels were determined with a Fuji

phosphorimager and plotted as the percentages of the values obtained with untreated cells.

Figures 12A and 12B show the dsRNA response in parental Huh7 and HeLa cells and HCV replicon-containing GS4.1 and SL1 cells. (12A) Phosphorylation of eIF-2 α . Huh7, GS4.1, HeLa, and SL1 cells were left untreated (lanes 1, 5, 9, and 13) or treated with 100 IU of IFN- α /ml for 12 h (lanes 2, 4, 6, 8, 10, 12, 14, and 16) and then transfected with poly(I:C) and incubated for 3 h (lanes 3, 4, 7, 8, 11, 12, 15, and 16). eIF-2 α -P and total eIF-2 α were determined by Western blots analysis with a monoclonal antibody specific for eIF-2 α -P and an antibody specific for total eIF-2 α protein. (12B) Induction of IFN- β mRNA by dsRNA and IFN- α . Parental Huh7 and HeLa cells and HCV replicon-containing GS4.1 and SL1 cells were left untreated (lanes 1, 5, 9, and 13) or treated with 100 IU of IFN- α /ml (lanes 3, 4, 7, 8, 11, 12, 15, and 16) for 12 h and then transfected with poly(I:C) (lanes 2, 4, 6, 8, 10, 12, 14, and 16) for 3 h. An RNase protection assay was performed with probes specific for IFN- β and β -actin mRNAs.

Figures 13A-13C show dose-dependent inhibition of the IFN- α response against subgenomes by lactacystin and epoxomicin. (13A) Cells were incubated with lactacystin and epoxomicin at the indicated concentrations for 7 h and then for an additional 12 h without the drugs. One hour after incubation with the proteasome inhibitors, IFN- α (100 IU/ml) was added for 6 h to a fraction of the cell culture plates (lanes 6 to 10 and 16 to 20). Viral RNA levels were determined by Northern blot analysis. rRNA was used as a control for the amount of RNA present in the samples. (13B and 13C) The amount of viral RNA was measured with a phosphorimager, and values were plotted as percentages of the values obtained with untreated cells.

Figures 14A and 14B are Northern blots and a graph showing that proteasome inhibitors block the IFN- α response against HCV replicons. (14A) GS4.1 cells were left untreated (lanes 1 to 3 and 19 to 21) or treated with 100 IU of IFN- α /ml for 6 h (lanes 4 to 6 and 22 to 24), with 5 μ M lactacystin (lanes 7 to 9 and 25 to 27) or 1 μ M epoxomicin (lanes 13 to 15 and 31 to 33) alone for 7 h, or with 5 μ M lactacystin (lanes 10 to 12 and 28 to 30) or 1 μ M epoxomicin (lanes 16 to 18 and 34 to 36) alone for 1 h and then in the presence of 100 IU of IFN- α /ml for an additional 6 h. Cells were harvested at 12 h (lanes 1 to 18) and 18 h (lanes 19 to 36) after addition of the cytokine. Viral RNA levels were determined by Northern blot analysis. rRNA was used as a control for the amount of RNA present in the samples. (14B) The amount of viral RNA was measured with a phosphorimager and the mean values and standard deviations from three samples were plotted. *, P < 0.05; **, P < 0.01. PSL, arbitrary units.

Figures 15A and 15B show a Northern blot graph demonstrating that proteasome inhibitors prevent establishment of an IFN- α response against HCV replicons. (15A) GS4.1 cells were treated with IFN- α for 10 h, followed by treatment with the indicated proteasome inhibitors for 12 h. Cells were left untreated (lanes 1 to 3 and 4 to 6) or treated with 100 IU of IFN- α /ml for 10 h (lanes 7 to 18), followed by treatment with 10 μ M lactacystin (lanes 13 to 15) or 1 μ M epoxomicin (lanes 16 to 18) for 12 h. Cells were harvested at 0, 10, and 18 h after the cytokine treatment, as indicated. Viral RNA levels were determined by Northern blot analysis. rRNA was used as a control for the amount of RNA present in the samples. (15B) The amount of viral RNA was measured with a phosphorimager and the mean values and standard deviations from three samples were plotted. PSL,

arbitrary units.

DETAILED DESCRIPTION OF THE INVENTION

The hepatitis C virus (HCV) pandemic affects the
5 health of more than 170 million people and is the major
indication for orthotopic liver transplantations (OLT).
Although the human liver is the primary site for HCV
replication, it is not known whether extrahepatic tissues
are also infected by the virus and whether non-primate
10 cells are permissive for RNA replication. However,
because viral replication leads to the accumulation of
mutations, it is conceivable that variants can emerge
with novel properties such as the potential to replicate
in different cell types of various species. Furthermore,
15 accumulation of a large number of quasispecies may also
contribute to resistance to IFN- α treatment. Therefore,
it is important to determine the properties of HCV
variants, and the effect such variation has on the
efficacy of IFN- α therapy.

20 Provided herein is evidence that subgenomic HCV RNAs
can replicate in mouse hepatoma and non-hepatic human
epithelial cells. Moreover, efficient replication
requires adaptation of the virus to cell-type specific
environmental conditions. These results show that HCV
25 RNA replication can lead to the accumulation of mutants
with altered tissue and host tropism thereby facilitating
the development of small animal models for HCV infection.

In accordance with the present invention, there are
provided nucleic acids and stably-transfected human non-
30 hepatic, and murine hepatic cell lines that replicate
HCV. Also provided are methods of use for such cells for
identifying therapeutic anti-viral agents for the
treatment of HCV infection. Additionally, the
availability of a murine line which replicates HCV
35 enables the production of a greatly needed mouse model of

HCV infection. Furthermore, the invention provides polynucleotides and their corresponding polypeptides which have adaptive mutations which result in expanded tropism of HCV.

5 The detailed description set forth below describes preferred methods for making and using the nucleic acids and cell lines of the present invention, and for practicing the methods of the invention. Any molecular cloning or recombinant DNA techniques not specifically
10 described are carried out by standard methods, as generally set forth, for example, in Sambrook et al., "DNA Cloning, A Laboratory Manual," Cold Spring Harbor Laboratory, 1989 and Ausubel et al. Current Protocols in Molecular Biology, J. Wiley & Sons, 1995.

15

I. Definitions

The following definitions are provided to aid in understanding the subject matter regarded as the invention.

20 As used herein, "hepatitis C virus" or "HCV" shall mean any representative of a diverse group of related viruses classified within the hepacivirus genus of the Flaviviridae family.

25 "Anti-HCV compounds" may include any inhibitor of HCV-derived enzymes, such as protease, helicase, and polymerase inhibitors. Anti-HCV compounds also include IRES inhibitors, glycosylation inhibitors, and molecules which block the HCV receptor (thus preventing entry into cells.) Other anti-HCV compounds include compounds which
30 enhance the specific or non-specific immune response, thereby ameliorating HCV infection or symptoms.

35 "HCV replication levels" may be measured by methods known in the art, including but not limited to detection of replicated HCV replicons, HCV NS protein production, or incorporation of detectably labeled nucleotides into

an HCV nucleic acid.

"Nucleic acid" or a "nucleic acid molecule" as used herein refers to any DNA or RNA molecule, either single or double stranded and, if single stranded, the molecule 5 of its complementary sequence in either linear or circular form. In discussing nucleic acid molecules, a sequence or structure of a particular nucleic acid molecule may be described herein according to the normal convention of providing the sequence in the 5' to 3' 10 direction. With reference to nucleic acids of the invention, the term "isolated nucleic acid" is sometimes used. This term, when applied to DNA, refers to a DNA molecule that is separated from sequences with which it is immediately contiguous in the naturally occurring 15 genome of the organism or virus in which it originated. When applied to RNA, the term "isolated nucleic acid" refers primarily to an RNA molecule that has been sufficiently separated from other nucleic acids with which it would be associated in its natural state (i.e., 20 in cells or tissues), and explicitly includes viral RNA. An isolated nucleic acid (either DNA or RNA) may further represent a molecule produced directly by biological or synthetic means and separated from other components present during its production.

25 "RNA subgenome" refers to any molecule which lacks some portion of a genome. For example, an RNA subgenome can be an HCV RNA molecule in which a structural gene has been replaced with a selection agent.

30 All amino acid residue sequences represented herein conform to the conventional left-to-right amino-terminus to carboxy-terminus orientation.

35 The term "isolated protein" or "isolated and purified protein" is sometimes used herein. This term refers primarily to a protein produced by expression of an isolated nucleic acid molecule of the invention.

Alternatively, this term may refer to a protein that has been sufficiently separated from other proteins with which it would naturally be associated, so as to exist in "substantially pure" form. "Isolated" is not meant to 5 exclude artificial or synthetic mixtures with other compounds or materials, or the presence of impurities that do not interfere with the fundamental activity, and that may be present, for example, due to incomplete purification, addition of stabilizers, or compounding 10 into, for example, immunogenic preparations or pharmaceutically acceptable preparations.

"Variants", "mutants" and "derivatives" of particular sequences of nucleic acids refer to nucleic acid sequences that are closely related to a particular 15 sequence but which may possess, either naturally or by design, changes in sequence or structure. By closely related, it is meant that at least about 75%, but often, more than 90%, of the nucleotides of the sequence match over the defined length of the nucleic acid sequence. 20 Changes or differences in nucleotide sequence between closely related nucleic acid sequences may represent nucleotide changes in the sequence that arise during the course of normal replication or duplication in nature of the particular nucleic acid sequence. Other changes may 25 be specifically designed and introduced into the sequence for specific purposes, such as to expand the tropism of viral RNA, or to change an amino acid codon or sequence in a regulatory region of the nucleic acid. Such specific changes may be made in vitro using a variety of 30 mutagenesis techniques or produced in a host organism placed under particular selection conditions that induce or select for the changes. Such sequence variants generated specifically may be referred to as "mutants" or "derivatives" of the original sequence. The terms 35 "percent similarity", "percent identity" and "percent

homology" when referring to a particular sequence are used as set forth in the University of Wisconsin GCG software program.

An "adaptive mutation" is a mutation in a nucleic acid sequence which produces a change in viral properties or activity. For example, an adaptive mutation includes, but is not limited to, a mutation which provides enhanced tropism for HCV, or which alters the efficacy of IFN- α treatment.

10 An HCV peptide, polypeptide, or protein of the invention includes any analogue, fragment, derivative or mutant which is derived from a HCV peptide or polypeptide and which retains at least one property or other characteristic of the HCV polypeptide. Different
15 "variants" of the HCV polypeptide exist in nature. These variants may be alleles characterized by differences in the nucleotide sequences of the gene coding for the protein, or may involve different RNA processing or post-translational modifications. The skilled person can
20 produce variants having single or multiple amino acid substitutions, deletions, additions or replacements. These variants may include inter alia: (a) variants in which one or more amino acids residues are substituted with conservative or non-conservative amino acids, (b)
25 variants in which one or more amino acids are added to the HCV peptide or polypeptide, (c) variants in which one or more amino acids include a substituent group, and (d) variants in which one or more amino acids are deleted from the HCV peptide or polypeptide. Other HCV peptides
30 or polypeptides of the invention include variants in which amino acid residues from one species are substituted for the corresponding residue in another species, either at the conserved or non-conserved positions. In another embodiment, amino acid residues at
35 non-conserved positions are substituted with conservative

or non-conservative residues. The techniques for obtaining these variants, including genetic (suppressions, deletions, mutations, etc.), chemical, and enzymatic techniques are known to the person having 5 ordinary skill in the art.

To the extent such variations, analogues, fragments, derivatives, mutants, and modifications, including alternative nucleic acid processing forms and alternative post-translational modification forms result in 10 derivatives of the HCV peptide or polypeptide that retain any of the biological properties of the HCV peptide or polypeptide, they are included within the scope of this invention.

The term "functional" as used herein implies that 15 the nucleic or amino acid sequence is functional for the recited assay or purpose.

The phrase "consisting essentially of" when referring to a particular nucleotide or amino acid means a sequence having the properties of a given sequence. 20 For example, when used in reference to an amino acid sequence, the phrase includes the sequence per se and molecular modifications that would not affect the fundamental and novel characteristics of the sequence.

A "replicon" is any genetic element, for example, a 25 plasmid, cosmid, bacmid, phage or virus, that is capable of replication largely under its own control. A replicon may be either RNA or DNA and may be single or double stranded.

A "vector", is a replicon, such as a plasmid, cosmid, 30 bacmid, phage or virus, to which another genetic sequence or element (either DNA or RNA) may be attached so as to bring about the replication of the attached sequence or element.

The phrase "operably linked" when referring to 35 nucleic acid constructs is used herein to indicate that

the respective promoter, operator and coding sequences, as well as any other 5' and 3' regulatory sequences, are arranged in the appropriate location, order and reading frame such that the desired control (e.g., expression) is
5 effected under appropriate conditions.

The term "oligonucleotide," as used herein refers to primers and probes of the present invention, and is defined as a nucleic acid molecule comprised of two or more ribo- or deoxyribonucleotides, preferably more than
10 three. The exact size of the oligonucleotide will depend on various factors and on the particular application and use of the oligonucleotide.

The term "probe" as used herein refers to an oligonucleotide, polynucleotide or nucleic acid, either
15 RNA or DNA, whether occurring naturally as in a purified restriction enzyme digest or produced synthetically, which is capable of annealing with or specifically hybridizing to a nucleic acid with sequences complementary to the probe. A probe may be either
20 single-stranded or double-stranded. The exact length of the probe will depend upon many factors, including temperature, source of probe and use of the method. For example, for diagnostic applications, depending on the complexity of the target sequence, the oligonucleotide
25 probe typically contains 15-25 or more nucleotides, although it may contain fewer nucleotides. The probes herein are selected to be "substantially" complementary to different strands of a particular target nucleic acid sequence. This means that the probes must be
30 sufficiently complementary so as to be able to "specifically hybridize" or anneal with their respective target strands under a set of pre-determined conditions. Therefore, the probe sequence need not reflect the exact complementary sequence of the target. For example, a
35 non-complementary nucleotide fragment may be attached to

the 5' or 3' end of the probe, with the remainder of the probe sequence being complementary to the target strand. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided 5 that the probe sequence has sufficient complementarity with the sequence of the target nucleic acid to anneal therewith specifically.

The term "specifically hybridize" refers to the association between two single-stranded nucleic acid 10 molecules of sufficiently complementary sequence to permit such hybridization under pre-determined conditions generally used in the art (sometimes termed "substantially complementary"). In particular, the term refers to hybridization of an oligonucleotide with a 15 substantially complementary sequence contained within a single-stranded DNA or RNA molecule of the invention, to the substantial exclusion of hybridization of the oligonucleotide with single-stranded nucleic acids of non-complementary sequence. For example, hybridizations 20 may be performed, according to the method of Sambrook et al., Molecular Cloning, Cold Spring Harbor Laboratory (1989), using a hybridization solution comprising: 5X SSC, 5X Denhardt's reagent, 1.0% SDS, 100 µg/ml denatured, fragmented salmon sperm DNA, 0.05% sodium 25 pyrophosphate and up to 50% formamide. Hybridization is carried out at 37-42°C for at least six hours. Following hybridization, filters are washed as follows: (1) 5 minutes at room temperature in 2X SSC and 1% SDS; (2) 15 minutes at room temperature in 2X SSC and 0.1% SDS; (3) 30 minutes-1 hour at 37°C in 1X SSC and 1% SDS; (4) 2 hours at 42-65°C in 1X SSC and 1% SDS, changing the 30 solution every 30 minutes.

One common formula for calculating the stringency conditions required to achieve hybridization between 35 nucleic acid molecules of a specified sequence homology

(Sambrook et al., 1989) is as follows:

$$T_m = 81.5^\circ\text{C} + 16.6 \log [\text{Na}^+] + 0.41(\% \text{ G+C}) - 0.63 (\% \text{ formamide}) - 600/\# \text{bp in duplex}$$

5

As an illustration of the above formula, using $[\text{Na}^+] = [0.368]$ and 50% formamide, with GC content of 42% and an average probe size of 200 bases, the T_m is 57°C . The T_m of a DNA duplex decreases by $1 - 1.5^\circ\text{C}$ with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of 42°C .

The stringency of the hybridization and wash depend primarily on the salt concentration and temperature of 15 the solutions. In general, to maximize the rate of annealing of the probe with its target, the hybridization is usually carried out at salt and temperature conditions that are $20-25^\circ\text{C}$ below the calculated T_m of the hybrid. Wash conditions should be as stringent as possible for 20 the degree of identity of the probe for the target. In general, wash conditions are selected to be approximately $12-20^\circ\text{C}$ below the T_m of the hybrid. In regards to the nucleic acids of the current invention, a moderate stringency hybridization is defined as hybridization in 25 $6X \text{ SSC}$, $5X$ Denhardt's solution, 0.5% SDS and $100 \mu\text{g}/\text{ml}$ denatured salmon sperm DNA at 42°C , and washed in $2X \text{ SSC}$ and 0.5% SDS at 55°C for 15 minutes. A high stringency hybridization is defined as hybridization in $6X \text{ SSC}$, $5X$ Denhardt's solution, 0.5% SDS and $100 \mu\text{g}/\text{ml}$ denatured 30 salmon sperm DNA at 42°C , and washed in $1X \text{ SSC}$ and 0.5% SDS at 65°C for 15 minutes. A very high stringency hybridization is defined as hybridization in $6X \text{ SSC}$, $5X$ Denhardt's solution, 0.5% SDS and $100 \mu\text{g}/\text{ml}$ denatured salmon sperm DNA at 42°C , and washed in $0.1X \text{ SSC}$ and 0.5% SDS at 65°C for 15 minutes.

The term "primer" as used herein refers to an oligonucleotide, either RNA or DNA, either single-stranded or double-stranded, either derived from a biological system, generated by restriction enzyme digestion, or produced synthetically which, when placed in the proper environment, is able to functionally act as an initiator of template-dependent nucleic acid synthesis. When presented with an appropriate nucleic acid template, suitable nucleoside triphosphate precursors of nucleic acids, a polymerase enzyme, suitable cofactors and conditions such as a suitable temperature and pH, the primer may be extended at its 3' terminus by the addition of nucleotides by the action of a polymerase or similar activity to yield a primer extension product. The primer may vary in length depending on the particular conditions and requirement of the application. For example, in diagnostic applications, the oligonucleotide primer is typically 15-25 or more nucleotides in length. The primer must be of sufficient complementarity to the desired template to prime the synthesis of the desired extension product, that is, to be able anneal with the desired template strand in a manner sufficient to provide the 3' hydroxyl moiety of the primer in appropriate juxtaposition for use in the initiation of synthesis by a polymerase or similar enzyme. It is not required that the primer sequence represent an exact complement of the desired template. For example, a non-complementary nucleotide sequence may be attached to the 5' end of an otherwise complementary primer. Alternatively, non-complementary bases may be interspersed within the oligonucleotide primer sequence, provided that the primer sequence has sufficient complementarity with the sequence of the desired template strand to functionally provide a template-primer complex for the synthesis of the extension product.

As used herein, the terms "reporter," "reporter system", "reporter gene," or "reporter gene product" shall mean an operative genetic system in which a nucleic acid comprises a gene that encodes a product that when expressed produces a reporter signal that is a readily measurable, e.g., by biological assay, immunoassay, radioimmunoassay, or by colorimetric, fluorogenic, chemiluminescent or other methods. The nucleic acid may be either RNA or DNA, linear or circular, single or double stranded, antisense or sense polarity, and is operatively linked to the necessary control elements for the expression of the reporter gene product. The required control elements will vary according to the nature of the reporter system and whether the reporter gene is in the form of DNA or RNA, but may include, but not be limited to, such elements as promoters, enhancers, translational control sequences, poly A addition signals, transcriptional termination signals and the like.

The terms "transform", "transfect", "transduce", shall refer to any method or means by which a nucleic acid is introduced into a cell or host organism and may be used interchangeably to convey the same meaning. Such methods include, but are not limited to, electroporation, microinjection, PEG-fusion and the like.

The introduced nucleic acid may or may not be integrated (covalently linked) into nucleic acid of the recipient cell or organism. In bacterial, yeast, plant and mammalian cells, for example, the introduced nucleic acid may be maintained as an episomal element or independent replicon such as a plasmid. Alternatively, the introduced nucleic acid may become integrated into the nucleic acid of the recipient cell or organism and be stably maintained in that cell or organism and further passed on or inherited to progeny cells or organisms of the recipient cell or organism. In other applications,

the introduced nucleic acid may exist in the recipient cell or host organism only transiently.

A "clone" or "clonal cell population" is a population of cells derived from a single cell or common 5 ancestor by mitosis.

A "cell line" is a clone of a primary cell or cell population that is capable of stable growth in vitro for many generations.

A "selectable marker" or a "selection agent" refers 10 to a nucleic acid sequence that when expressed confers a selectable phenotype, such as antibiotic resistance, to a transformed cell.

A "viral antigen" shall be any peptide, polypeptide or protein sequence, segment or epitope that is derived 15 from a virus that has the potential to cause a functioning immune system of a host to react to said viral antigen.

An "antibody" or "antibody molecule" is any immunoglobulin, including antibodies and fragments 20 thereof, that binds to a specific antigen. The term includes polyclonal, monoclonal, chimeric, and bispecific antibodies. As used herein, antibody or antibody molecule contemplates both an intact immunoglobulin molecule and an immunologically active portion of an immunoglobulin 25 molecule such as those portions known in the art as Fab, Fab', F(ab')₂ and F(v).

The term "detectably label" is used herein to refer to any substance whose detection or measurement, either directly or indirectly, by physical or chemical means, is 30 indicative of the presence of the target bioentity in the test sample. Representative examples of useful detectable labels, include, but are not limited to the following: molecules or ions directly or indirectly detectable based on light absorbance, fluorescence, 35 reflectance, light scatter, phosphorescence, or

- luminescence properties; molecules or ions detectable by their radioactive properties; molecules or ions detectable by their nuclear magnetic resonance or paramagnetic properties. Included among the group of 5 molecules indirectly detectable based on light absorbance or fluorescence, for example, are various enzymes which cause appropriate substrates to convert, e.g., from non-light absorbing to light absorbing molecules, or from non-fluorescent to fluorescent molecules.
- 10 As used herein, the term "living host" shall mean any non-human autonomous being.

II. Methods for Obtaining HCV RNA and Producing Non-Hepatic Human Cell Lines and Non-Human Hepatic Cell Lines

15 that Replicate HCV.

The HCV replicating non-hepatic human cell-based and non-human hepatic cell-based systems are prepared according to the general methods set forth below for isolation of nucleic acids, transformation of cultured 20 cells, and maintenance of cell lines.

A. Nucleic acids

The HCV replicons of the present invention comprise adaptive mutations which alter the ability of HCV to 25 replicate in different cell types. Surprisingly, the present inventors have identified mutations which are associated with expanded viral tropism.

The HCV nucleic acid molecules of the invention may be prepared by two general methods: (1) They may be 30 synthesized from appropriate chemical starting materials, or (2) they may be isolated from biological sources. Both methods utilize protocols well known in the art.

The availability of nucleotide sequence information enables preparation of an isolated nucleic acid molecule 35 of the invention by oligonucleotide synthesis. Synthetic

oligonucleotides may be prepared by the phosphoramidite method employed in the Applied Biosystems 38A DNA Synthesizer or similar devices. The resultant construct may be purified according to methods known in the art, 5 such as high performance liquid chromatography (HPLC). Long, double-stranded polynucleotides, such as a DNA molecule of the present invention, must be synthesized in stages due to the size limitations inherent in current oligonucleotide synthetic methods. Thus, for example, a 10 3 kilobase double-stranded molecule may be synthesized as several smaller segments of appropriate complementarity. Complementary segments thus produced may be ligated such that each segment possesses appropriate cohesive termini for attachment of an adjacent segment. Adjacent segments 15 may be ligated by annealing cohesive termini in the presence of DNA ligase to construct an entire 3 kilobase double-stranded molecule. A synthetic DNA molecule so constructed may then be cloned and amplified in an appropriate vector.

20 HCV nucleic acid sequences may be isolated from appropriate biological sources using methods known in the art. For example, total RNA can be extracted with TRIzol reagent from Gibco BRL, although other reagents are also available for this purpose.

25 In some cases, it may be desirable to synthesize HCV subgenomic RNA wherein a selectable marker gene is substituted for a HCV structural gene.

The availability of HCV replicon encoding nucleic acids enables the production of strains of laboratory 30 mice carrying part or all of the HCV sequence or mutated sequences thereof. Such mice provide an *in vivo* model for studying HCV infection, and analyzing possible treatment modalities for the same.

Methods of introducing transgenes in laboratory mice 35 are known to those of skill in the art. Three common

methods include: 1. integration of retroviral vectors encoding the foreign gene of interest into an early embryo; 2. injection of DNA into the pronucleus of a newly fertilized egg; and 3. the incorporation of 5 genetically manipulated embryonic stem cells into an early embryo.

A transgenic mouse carrying an HCV replicon comprising the adaptive mutations is generated by genomic integration of exogenous genomic sequence encoding HCV. 10 These transgenic animals are useful for drug screening studies as animal models for human diseases.

The term "animal" is used herein to include all vertebrate animals, except humans. It also includes an individual animal in all stages of development, including 15 embryonic and fetal stages. A "transgenic animal" is any animal containing one or more cells bearing genetic information altered or received, directly or indirectly, by deliberate genetic manipulation at the subcellular level, such as by targeted recombination or 20 microinjection or infection with recombinant virus. The term "transgenic animal" is not meant to encompass classical cross-breeding or *in vitro* fertilization, but rather is meant to encompass animals in which one or more cells are altered by or receive a recombinant DNA 25 molecule. This molecule may be specifically targeted to a defined genetic locus, be randomly integrated within a chromosome, or it may be extrachromosomally replicating DNA. The term "germ cell line transgenic animal" refers to a transgenic animal in which the genetic alteration or 30 genetic information was introduced into a germ line cell, thereby conferring the ability to transfer the genetic information to offspring. If such offspring, in fact, possess some or all of that alteration or genetic information, then they, too, are transgenic animals.

35 A type of target cell for transgene introduction is the embryonal stem cell (ES). ES cells may be obtained from pre-implantation embryos cultured *in vitro* (Evans et

al., (1981) *Nature* **292**:154-156; Bradley et al., (1984) *Nature* **309**:255-258; Gossler et al., (1986) *Proc. Natl. Acad. Sci.* **83**:9065-9069). Transgenes can be efficiently introduced into the ES cells by standard techniques such
5 as DNA transfection or by retrovirus-mediated transduction. The resultant transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The introduced ES cells thereafter colonize the embryo and contribute to the germ line of the resulting
10 chimeric animal.

B. Cell lines

The cell lines of the invention include any cell which supports production of HCV components. These cells
15 include human, non hepatic cells and/or non-human hepatic cells, such as murine hepatic cells. Cell lines useful for practice of the invention include, but are not limited to HELA, a non-hepatic epithelial cell line (ATCC CRL number CCL-2.2), Hepal-6, a murine hepatoma cell line
20 (ATCC CRL number-1830), and AML-12, a murine hepatocyte cell line (ATCC CRL number-2254).

To achieve stable gene transfer, HCV subgenomic RNA is introduced into host cells. This may be accomplished according to numerous methods known in the art,
25 including, but not limited to: (1) calcium phosphate transfection; (2) transfection with DEAE-dextran; (3) electroporation; and (4) liposome-mediated transfection. For general protocols, see, e.g., chapter 9 in *Current Protocols in Molecular Biology*, Ausubel et al. (editors),
30 John Wiley & Sons, Inc. 1987-1995. For stable transfer of nucleic acids into mammalian cells, the liposome-mediated transfection method may be used in the present invention because of the large amount of nucleic acid that can be introduced into the cells, thereby increasing the
35 possibility of integration of the nucleic acid into the host genome.

Cells are grown according to standard methods known in the art, such as those set forth in Culture of Animal Cells: A Manual of Basic Technique by R. Ian Freshney, 4th ed. Edition, available from the ATCC.

- 5 Stable transfectants are selected by the ability of an individual cell colony to grow in the presence of a selection agent, e.g., an antibiotic, by virtue of a resistance-encoding gene on HCV RNA or by isolating cells using FACS and antibodies directed any HCV protein.
- 10 Detection and quantitation of expression of HCV gene products in stably-transfected cell lines of the invention can be accomplished using a variety of known assays. For instance, cells transformed with the RNA subgenomes of HCV can be selected with an antibiotic such
- 15 as G418 (neomycin). Alternatively, cells may be selected based on the presence and accumulation of HCV RNA or HCV gene products. As another example, the starting HCV encoding nucleic acids may be modified to also comprise a hygromycin or puromycin or any other resistance or
- 20 reporter gene, such that cells transfected with the nucleic acids can be selected by their ability to grow on hygromycin- or puromycin-containing medium. Alternatively, selectable markers including luciferase, beta lactamase etc., may be utilized which allow for the
- 25 selection of cells by FACS and related procedures. In an alternative embodiment, a separate plasmid may be constructed that comprises an antibiotic resistance gene, and can be used to co-transfect cells along with the subgenomic RNA molecules. Further, as described in
- 30 detail in the following Examples, cells stably transfected with the subgenomic HCV RNA are grown in the appropriate medium for a selected period of time, the medium is then collected and analyzed for the presence of HCV RNA by dot blot hybridization or by conventional Northern hybridization, using a radioactively labeled
- 35

probe having HCV DNA or RNA complementary sequences. Alternatively, viral gene products may be detected in the cells of the invention using conventional methods, including, without limitation, immunoassay and Western blotting.

Using the assays described above, stably-transfected cell lines can be selected which possess optimum characteristics for use in cell-based assays for screening potential anti-viral compounds.

Another aspect of the invention includes a non-human host animal which comprises the HCV expressing cells of the invention. These animals may be produced by administration of a HCV replicating cell, an HCV encoding nucleic acid having one or more adaptive mutations which permit replication in mice. The cells or viral nucleic acid could be directly injected intravenously (e.g. via tail vein injection), intramuscularly, subcutaneously, or via-intrahepatic injection. Alternatively, transgenic mice could be produced using the HCV replicons of the inventions, as described above.

III. Uses of Cell Lines for Cell-Based Assays of Potential Anti-HCV Agents

The human non-hepatic and murine hepatic cell lines of the invention which replicate HCV may be used in research, diagnostic, and therapeutic applications, including cell-based assays to evaluate the effectiveness of potential anti-HCV compounds, utilizing methodologies known in the art. Typical assays are summarized herein below. These cell-based assays may be performed in standard cell culture media utilizing commonly-available equipment, reagents and culture containers.

Persons skilled in the art will appreciate that these assays represent exemplary embodiments, and may be varied to provide similar/equivalent equipment or

reaction conditions. For example, a variety of genes encoding antibiotic resistance are available, and can be utilized in accordance with the present invention in the generation of the cell lines of the invention. In a
5 preferred embodiment, RNA isolated from parental human hepatic or untransformed cells is also utilized as a control in the assays described herein below to determine the effects of potential anti-viral compounds on HCV expressed in the cells. The control RNA is obtained in a
10 manner similar to the HCV RNA. This cell line is treated in the assays described herein below as a negative control, to assure that any effects observed are due to the action of the compound being tested on HCV, and not non-specific effects due to the introduction of RNA into
15 the cells.

A. General Cell-Based Assay for Inhibitors of HCV replication

96-well microtiter plates are seeded with an
20 appropriate amount of cells which replicate HCV in a standard cell culture medium containing G418 (e.g., 400 µg/ml), as well as standard concentrations of penicillin, streptomycin and kanamycin or gentamicin to prevent bacterial and mycoplasma contamination. The cells are
25 incubated at 37°C in a humidified 5% CO₂ incubator. On day 0 wells are washed three times with warm phosphate-buffered saline (PBS). The culture medium is then replaced with fresh medium containing 0.3% dimethylsulfoxide (DMSO), 10% fetal calf serum (FCS),
30 penicillin, streptomycin, kanamycin/gentamicin, containing one of the following ingredients: (1) various concentrations of a known HCV inhibitor, such as interferon alpha, as a positive control; and (2) various concentrations of one or more of the compounds to be
35 tested. The plates are incubated at 37°C in humidified,

5 5% CO₂ incubator for 24, 48, and 72 hours. The plates
are washed twice with PBS and then with a solution of
methanol and acetone (1:1) to fix the cells. The cells
are then incubated with an antibody specific for a viral
protein (i.e. NS5A) according to the standard methods,
such as enzyme linked immunosorbent assay (ELISA).
Briefly, following incubation with the primary antibody,
the plates are washed to remove unbound antibody and then
10 incubated with a second, enzyme-conjugated antibody that
can bind to the primary antibody. The plates are washed
again, followed by an incubation with a colorless
substrate that upon hydrolysis (cleavage) by the enzyme
yields a colored product, the concentration of which can
be determined with a spectrophotometer (microtiter plate
15 reader). The concentration of the product corresponds to
the levels of viral replication in cells and can be used
to determine the activity of a given drug to inhibit HCV
replication.

20 **B. Cytotoxicity Assays**

A cytotoxicity assay may be conducted to evaluate
potential anti-HCV agents, utilizing a protocol similar
to that described above. Instead of measuring HCV
replication levels, however, cytotoxicity of the various
25 test agents is assessed by standard procedures to
determine cell viability, proliferation and levels of
cellular metabolism including but not restricted to cell
membrane permeability, lysosomal mass-pH, cell density or
mitochondrial activity. For example, the CytoTox-ONE™
30 Assay from Promega is a rapid, fluorescent measure of the
release of lactate dehydrogenase (LDH) from cells with a
damaged membrane. LDH released into the culture medium
is measured with a 10-minute coupled enzymatic assay that
results in the conversion of resazurin into resorufin.
35 Since the CytoTox-ONE™ Reagent mix does not damage

healthy cells, released LDH can be measured directly in assay wells containing a mixed population of viable and damaged cells.

5 IV. Identification of Cell Lines Permissive for HCV Infection

As shown herein, it is possible to produce HCV carrying adaptive mutations that confer broad tissue and species tropism. Using such virus stocks it will be
10 possible to screen cell lines of human and non-human origin for virus infection. Briefly, probes which correspond to unique portions of the sequence, may be used in detection methods. This method will lead to the identification of novel cell lines that are permissive
15 for a complete cycle of HCV replication.

V. Screening for the HCV Receptor(s)

The ability to replicate HCV in different cell lines facilitates the isolation of the HCV receptor. Virus stocks similar to the ones described in section IV can be used to isolate the HCV receptor(s). For this purpose virus stocks carrying replicons with a selectable marker, such as neomycin or hygromycin will be used. Cells that are non-permissive for infection, will be transfected
25 with DNA isolated from cells that are known to express the receptor (i.e. human hepatocytes, cells identified with the procedure described in section III) and subsequently infected with recombinant HCV carrying the selectable marker. Cells that express the receptor can
30 then be selected through the addition of an antibiotic (i.e. G418 or hygromycin) to the culture medium. Once cells are identified, the transfected DNA can be isolated, cloned, and sequenced. The sequence information can then be used to identify the gene(s)
35 encoded by transfected DNA.

The following examples are provided to describe the invention in further detail. These examples are intended to illustrate and not to limit the invention.

5

EXAMPLE I Human Non-Hepatic and Mouse Hepatic Cell Lines
that replicate HCV

MATERIALS AND METHODS

- 10 Cell culture. Cells were purchased from the American Type Culture Collection (BHK Kidney *Mesocricetus auratus* (Syrian golden hamster) ATCC CRL-1632; Vero Kidney epithelial *Cercopithecus aethiops* (African green monkey) ATCC CCL-81; CV-1 Kidney fibroblast *Cercopithecus*
- 15 *aethiops* (African green monkey) ATCC CCL-70; HT1080 Fibrosarcoma *Homo sapiens* (human) ATCC CRL12012; HeLa Cervix carcinoma *Homo sapiens* (human) ATCC CCL2; McA-RH7777 Hepatoma *Rattus norvegicus* (rat) ATCC CRL-1601; FTO2B Hepatoma *Rattus norvegicus* (rat); Hepal-6 Hepatoma
- 20 *Mus musculus* (mouse) ATCC CRL-1830; AML12 Hepatocyte *Mus musculus* (mouse) ATCC CRL-2254; FL83B Hepatocyte *Mus musculus* (mouse) ATCC CRL-2390). The Huh7-derived cell lines GS4.1 and GS4.5 are subclones derived from cell lines FC41 and FCA4, respectively (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). Cell line Bsp8 is a Huh7-derived cell line expressing HCV-N subgenomic replicon 1bneoΔS (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). All cultures were grown in Dulbecco's modified Eagle's medium (Gibco-Invitrogen) supplemented with 10% fetal bovine serum, L-glutamine, nonessential amino acids, penicillin, and streptomycin.
- 25
- 30

- 35 RNA transfection. All the plasmids were linearized with ScaI, and RNA was synthesized with the MEGAscript kit (Ambion). In vitro-transcribed RNA was purified as

previously described (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). Total cellular RNA was extracted with Trizol reagent (Invitrogen). The conditions used for the transfection of cells with total RNA were identical
5 to those used for the transfection with in vitro-transcribed RNA (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). Colonies were selected with G418 at a concentration of 1 mg/ml.

10 RNA analysis. Total cellular RNA was extracted from transfected cell lines with Trizol reagent. Five micrograms of total RNA was fractionated on 1% agarose gels containing 2.2 M formaldehyde and transferred onto a nylon membrane. Membranes were hybridized with riboprobes
15 specific for plus-stranded HCV replicon RNA, human papillomavirus (HPV) E6, and mouse albumin mRNA as described previously (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). The HPV and mouse albumin probes spanned nucleotides 811 to 1491 (GenBank accession number
20 M20325) and nucleotides 1501 to 1988 (GenBank accession number XM_132149), respectively.

Reverse transcription-PCR and DNA sequencing. Nucleotide and amino acid numbers correspond to the HCV type 1b genome Con-1 (AJ238799). HCV replicons were isolated and cloned from established cell lines by PCR amplification of three fragments spanning the entire NS region from position 3420 to 9410. The untranslated regions at the 5' and 3' ends of HCV RNA were cloned separately for
30 nucleotide sequence analysis. DNA synthesis was carried out with Superscript II reverse transcriptase provided in a cDNA synthesis kit (Gibco- Invitrogen). The DNA oligomers used as primers for the reverse transcription reaction mapped to positions 485 to 465, 5492 to 5473, 35 7256 to 7234, 9410 to 9388, and 9616 to 9597. The

reaction mixtures were incubated for 1 h at 45°C. PCR was performed with an Advantage PCR kit (Clontech). One microliter of the cDNA reaction mixture was used for PCRs with 19- to 23-nucleotide-long primers that yielded 5 fragments spanning positions 1 to 464, 1387E to 5082, 5016 to 7226, 7154 to 9387, and 9239 to 9616. Position 1387E refers to an oligomer specific for the encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) element located upstream of NS3. The PCR 10 products were cloned into plasmid pGEM-T Easy (Promega). Four clones of each fragment were sequenced with an ABI automatic DNA sequencer, and a consensus sequence was established with the help of a sequence assembly program (Genetics Computer Group).

15 Long reverse transcription-PCR was performed with an Advantage-GC kit (Clontech) with a pair of primers beginning at positions 1415E, upstream of NS3, and 7989 within NS5B. The PCR conditions were modified as follows: step 1, 95°C for 3 min; step 2, 5 cycles, 30 s at 95°C 20 and 6 min at 72°C; step 3, 27 cycles, 30 s at 95°C and 6 min at 68°C; step 4, 68°C for 6 min. PCR products were gel purified and digested with HindIII and MfeI and replaced with the corresponding fragment in plasmid I377/NS3-3'.

25 Plasmid construction. All plasmids (Table 3) were derived from the parental HCV Con-1 replicon I377/NS3-3' (AJ242652). Subgenomes containing consensus mutations were constructed by replacing DNA restriction fragments 30 with the corresponding fragments from the pGEM-T Easy cDNA libraries (see above). The resulting plasmids with the amino acid changes in the NS region are listed in Table 3.

35 Immunofluorescence. Cells were plated on coverslips in

six-well plates at least 16 h before treatment, washed with phosphate-buffered saline, and fixed with cold methanol-acetone (1:1) for 15 to 20 min. Next, the cells were blocked in phosphate-buffered saline containing 10%
5 fetal bovine serum for 30 min at room temperature and then incubated with anti-NS5A antibodies (a gift from Chen Liu) and fluorescein isothiocyanate-conjugated goat anti-mouse immunoglobulin antibodies (Jackson Laboratories). In addition, cells were stained with the
10 DNA binding fluorochrome DAPI (4',6'-diamidino-2-phenylindole). Coverslips were mounted with antifade agent (Molecular Probes), examined with a Nikon immunofluorescence microscope, and photographed with a charge-coupled device camera.

15

RESULTS

HCV replication in cells of nonhepatic origin.

HCV exhibits a very narrow host range and infects only humans and chimpanzees. We question whether this
20 limitation was due to determinants of RNA replication. Because efficient replication of subgenomes depends on genetic adaptations of the replicon (Blight, K. J., et al. 2000. Science 290:1972-1975; Guo, J. T., et al., 2001., J. Virol. 75:8516-8523; Lohmann, V., et al., 2001. J. Virol. 75:1437-1449), presumably to compensate for subtle variations in the cellular environments among cells from different tissues, it was hypothesized that replication in cells of nonhepatic origin would require additional, cell-type-specific adaptive mutations.
25 Transfection of several primate- and rodent-derived cell lines with subgenomic RNA transcribed from plasmid DNA carrying previously identified adaptive mutations in Huh7 cells did not yield cell lines expressing replicons. To increase the chance for the selection of RNA subgenomes
30 capable of replicating in cells of nonhepatic origin,
35

subgenomic RNA isolated from Huh7 cell lines that replicate HCV RNA was used. Because of the high rate of nucleotide incorporation errors that occur during RNA-directed RNA synthesis, this population of viral 5 subgenomes exhibited much greater genetic heterogeneity than did RNA transcribed from a DNA template in vitro.

Upon transfection of HeLa cells with total RNA obtained from Huh7 cell lines GS4.1, GS4.5, and Bsp8, G418-resistant cell clones were obtained. The number of 10 clones ranged from approximately 2 (Bsp8) to 50 (GS4.1) per 10 µg of total RNA depending on the origin of the RNA used for the transfections. Replicons in these three Huh7-derived cell lines contained different adaptive mutations and replicated two different HCV 1b genomes 15 (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523). Several HeLa-derived colonies obtained with total RNA from GS4.1 cells were subsequently expanded into seven stable cell lines (SL1 to SL7; Fig. 1A, lanes 4 and 8 to 12). The amounts of viral RNA present in early passages 20 of these cell lines examined ranged from 0.05 to 7.5 ng/10 µg of total RNA, which corresponded to 20 to 3,000 copies of RNA per cell. In general, the amounts of RNA increased upon passage of cells and reached levels that were comparable to those obtained with the most 25 productive Huh7- derived cell lines such as GS4.1 (lanes 2 and 4 to 6). As expected, expression of viral gene products could be confirmed by immunofluorescence with antibodies directed against NS5A (Fig. 1B). As with GS4.1 cells, more than 90% of SL1 cells expressed viral 30 proteins. However, in contrast to Huh7 cell lines where the accumulation of HCV RNA declines approximately 100-fold when cells become confluent, viral replication in HeLa cells was not affected by the growth conditions of the cells, i.e., SL1 cells continued to produce high 35 amounts of viral RNA even when they became confluent

(results not shown) (Guo, J. T., et al., 2001., J. Virol. 75:8516-8523; Pietschmann, T., et al. 2001. J. Virol. 75:1252-1264).

5 Adaptation of HCV replicons.

To determine whether HCV replication in HeLa cells led to the selection of subgenomes with cell-type-specific adaptive mutations, the efficiency by which G418-resistant colonies formed in Huh7 and HeLa cells 10 transfected with total RNA isolated from GS4.1 and SL1 cells was compared. Total RNA from GS4.1 cells led to the selection of approximately 166 G418-resistant colonies per ng of viral RNA in Huh7 cells compared with only 4 colonies in HeLa cells (Table 1). In contrast, total RNA 15 from SL1 cells yielded 160 colonies in HeLa cells compared with about 20 in Huh7 cells. These results indicated that replication in HeLa cells led to the selection of variants with cell-type-specific adaptive mutations that were responsible for the 40-fold increase 20 in colony formation efficiency between amplified RNA in GS4.1 and SL1 cells. Nucleotide sequence analysis of HCV cDNA clones obtained from the SL1 and SL2 cell lines confirmed this view. These data showed that replicons in the two HeLa cell lines maintained the previously 25 identified adaptive mutations in GS4.1 cells and acquired several additional mutations that resulted in amino acid changes in the NS region (Figure 2 and Table 2). Notably, some of the new mutations formed clusters in the NS4B and NS5A regions. In the case of SL1 cells; a deletion of 43 30 amino acids near the C terminus of NS5A was observed. Of particular interest were mutations in the amino-terminal region of NS4B, because they have so far not been found in cDNAs from replicons in Huh7 cells and hence could have been responsible for the observed adaptation of 35 replicating RNA (Blight, K. J., et al. 2000. Science

290:1972-1975; Guo, J. T., et al., 2001., J. Virol. 75:8516-8523; Krieger, N., et al., 2001. J. Virol. 75:4614-4624; Lohmann, V., et al., 2001. J. Virol. 75:1437-1449). Moreover, one mutation at position 1749 5 was present in both SL1 and SL2 cells. In contrast to the results obtained with the NS regions, no mutations were detected in the 5' and 3' untranslated regions of replicons expressed in SL1 and SL2 cells.

10

Table 1. Colony formation efficiency of total cellular RNA^a.

Cell	Viral RNA (ng/10 µg)	No. of colonies in transfected cells					
		HuH7		HeLa		Hepal-6	
		Mean (SD)	Colony/ ng of viral RNA	Mean (SD)	Colony/ ng of viral RNA	Mean (SD)	Colony/ ng of viral RNA
GS4.1	5	834 (64)	166	22 (4)	4	0	<1
SL1	5	100 (53)	20	803 (81)	160	1.3 (1.5)	<1
MH1	0.5	20 (2)	40	66 (9)	132	1.7 (0.6)	3

^a Results from three independent transfection experiments. Total RNA was extracted from GS4.1, SL1, and MH1 cells at passages 21, 26, and 4, respectively.

30

35

40

45

TABLE 2. Consensus mutations in replicons isolated from HeLa and mouse hepatoma cell clones

Cell clone	Conserved mutation(s)	NS protein
GS4.1	E1202G S2204I, D2254E, I2324V	NS3 NS5A
SL1	Q1067R, S1128A, E1202G, S1323P, S1560G ^a L1701F	NS3 NS4A
10	Q1720R, Q1727H, V1749A, V1893L T2035A, S2204I, I2252V, D2254E, I2274V, R2290L, I2324V, del.2371-2413 ^b	NS4B NS5A NS5B
15	W2990R SL2	NS3 NS4B
	I1097V, Q1112R, P1115L, V1593M, M1647I L1715P, Q1737R, V1749A, I1797V, N1965Y	NS5A
	Q2012L, S2204I, E2247G, D2254E, K2302R, I2324V, S2336P, L2400S, E2411G, A2412V	NS5A
MH1	Q1067R, S1128A, E1202G, S1323P, S1560G Q1720R, Q1727H, V1749A, V1893L	NS3 NS4B
20	T2035A, S2204I, I2252V, D2254E, I2274V, R2290L, I2324V, M2388T, T2496A W2990R	NS5A NS5B
MH2	Q1112R, E1202G, S1323P, S1560G L1701F	NS3 NS4A
25	Q1720R, Q1727H, V1749A, V1893L T2035A, T2185A, S2204I, I2252V, D2254E, I2274V, R2290 ^a , I2324V ^a W2990R ^a	NS4B NS5A NS5B
MH4	Q1067R, S1128A, E1202G, S1323P, S1560G L1701F	NS3 NS4A
30	Q1720R, Q1727H, V1749A, V1893L, A1841T T2035A, S2204I, I2252V, D2254E, I2274V, R2290L, I2324V, T2364M, L2391R I2843V, W2990R	NS4B NS5A NS5B

^a Mutation that occurred in 50% of the clones analyzed.

^b del., deletion.

Mouse hepatoma cells can support HCV RNA replication.

The discovery of several additional mutations in cDNA clones obtained from SL1 and SL2 cells suggested total RNA from these cell lines might yield colonies in cells that did not appear to be permissive for HCV replication after transfection with subgenomic RNA or total RNA from Huh7-derived cell lines. Hepatoma and hepatocyte-derived cell lines were examined. G418-

resistant colonies were obtained with the mouse hepatoma cell line Hepal-6 after transfection with total RNA from SL1 cells (Figure 3A, lanes 4 to 6, 9, and 12). As with HeLa cells, the amounts of RNA ranged from 300 to 1,000
5 copies of RNA per cell and a large fraction of the cells expressed viral proteins (Figure 3B). In contrast to Huh7 and HeLa cells, the amount of HCV RNA in the mouse cell lines appeared to vary between cell passages (Figure 3A, lanes 6 to 14). Interestingly, total RNA isolated from
10 one of the mouse cell lines, MH1, did not produce significantly more colonies in Hepal-6 cells than did total RNA from SL1 cells, suggesting that the subgenomes present in SL1 cells were already adapted for replication in the mouse cells (Table 1). In support of this
15 interpretation, nucleotide sequence analysis of viral cDNAs cloned from three mouse cell lines showed that the majority of the mutations identified in SL1 cells were maintained (Figure 2). Surprisingly, the deletion in NS5A identified in four of four clones sequenced from SL1
20 cells was not present in replicons isolated from mouse cells, indicating that a subpopulation of replicons without the deletion was still present in these (SL1) cells.

In further experiments, mouse hepatocyte cells AML12
25 (ATCC CRL-2254) were transfected with total RNA isolated from the cell line GS4.1, expressing subgenomic replicons and from cell line HFL expressing full-length HCV genomes, respectively. G418 resistant colonies were isolated to establish stable cell lines expressing HCV
30 subgenomic and full-length replicons. 5 micrograms of total RNA was isolated from AML12 cell lines (MA6-5 to MA6-8, and MAC1) that were established from G418 resistant cell colonies and analyzed by Northern blot analysis, which confirmed replication of HCV (Figure 4).

Cell-derived HCV RNA is more efficient than in vitro-transcribed RNA in initiating replication in HeLa and mouse hepatoma cells.

Results showed that replication of HCV subgenomes in 5 HeLa and mouse cells led to the selection of replicons with several novel mutations. The majority of these mutations were located in the NS3, NS4B, and NS5A regions. Moreover, the results showed that cell-derived RNA carrying some or all of these mutations was much more 10 efficient in establishing G418-resistant colonies in HeLa cells than was RNA derived from Huh7 cells (Table 1).

Based on these observations, it was surmised that introduction of these mutations into available subgenomic replicons should alter or expand their tissue and host 15 tropism. To test this hypothesis, 13 subgenomic replicons were designed that carried mutations in NS3, NS4B, and NS5A alone or in combination with each other as described in Table 3. Of the 13 constructs examined, only two, pZS2 and pZS25, yielded a small number of G418- resistant 20 colonies in HeLa cells (Table 4). Viral RNA replication was confirmed by Northern blot analysis of total RNA isolated from six cell lines derived from those colonies. None of the variants yielded colonies in Hepa1-6 cells. Moreover, negative-control experiments with in vitro- 25 transcribed RNA derived from a variant containing a frameshift mutation in NS5B did not yield any colonies that could be expanded into cell lines. Notably, save for one, all replicons were permissive for replication in Huh7 cells, albeit with significantly different efficiencies (Table 3). Interestingly, both pZS2 and pZS25 carried mutations in NS4B that were conserved in replicons from two independent HeLa cell lines, SL1 and SL2. In addition, these replicons had the S2204I mutation 30 in NS5A that was previously found to be one of the most potent adaptive mutations for HCV replication in Huh7 35

cells. Because both replicons replicated very efficiently in Huh7 cells, the results suggested that the NS4B mutations could have contributed to the observed expansion of the tissue tropism of HCV replicons. In support of this hypothesis, the subgenome with the highest efficiency in Huh7 cells, pZS11 lacking mutations in NS4B (Table 3), did not yield any colonies in HeLa cells.

10 TABLE 3. Colony formation efficiency of mutant replicons
in Huh7 cells

15 Vector ¹	Conserved mutation(s)	Mean (SD) CFE/ pg of RNA ²	NS protein(s) affected
I _{3'UTR} /NS3-3'	NA ^b	2.3 (1.5)	NA
pZS10	A1 ^a	3.3 (3.5)	NS3
pZS1	C1 ^b	1.4 X 10 ³ (3.9 X 10 ²)	NS5A
pZS20	B ^c	1.7 (0.6)	NS4AB
20 pZS11	A1 + C1	2.4 X 10 ⁵ (7 X 10 ⁴)	NS3, NS5A
pZS2	B + C1	7.8 X 10 ⁴ (1.7 X 10 ⁴)	NS4AB, NS5A
pZS12	A1 + B	0.3 (0.6)	NS3, NS4AB
pZS5	A1 + B + C1	15	NS3, NS4AB, NS5A
25 pZS4	B + C1 + A2 ^d	165 (21)	NS3, NS4AB, NS5A
pZS8	A2 + B	0	NS3, NS4AB
pZS6	A2 + B + C2 ^e	8 (4)	NS3, NS4AB, NS5A
pZS15	C3 ^f	491 (183)	NS5A
pZS25	B + C3	1.7 X 10 ⁴ (2.6 X 10 ³)	NS4AB, NS5A
30 pZS45	A2 + B + C3	19 (2)	NS3, NS4AB, NS5A

^a A1, mutation E1202G.

^b C1, mutations S2204I and D2254E.

^c B, mutations L1701F, Q1720R, Q1727H, and V1749A.

^d A2, mutations E1202G, S1128A, and S1323P.

^e C2, mutations S2204I, D2254E, R2290I, and I2324R.

^f C3, all the mutations of C2 plus the deletion 2371-2413.

²⁵ ² CFE, colony formation efficiency. Values are derived from three independent transfections of each replicon RNA.

^b NA, not applicable.

40 ¹ Sequence ID Numbers for subgenomic replicons are listed in Table 6

TABLE 4. Colony formation efficiency of in vitro-transcribed RNA^a

5	Source of cDNA library (cell line or plasmid)	No. of colonies in transfected cells		
		Huh7	HeLa	Hepal-6
10	GS4.1	>10,000	0, 0, 0	0, 1, 0
	SL1	>10,000	0, 3, 2	0, 1, 0
	MH4	>10,000	3, 4, 0	17, 0, 0
	pZS2	>10,000	2, 3, 0	0, 0, 0
	pZS25	>10,000	0, 2, 1	0, 0, 0

15 ^a Results from transfection experiments with in vitro-transcribed RNA from pooled clones isolated from the indicated cell line and from in vitro-transcribed RNA from pZS2 and pZS25 (Table 3).

20 To further explore the basis for the observed low colony formation efficiency of in vitro-transcribed RNA in HeLa cells, it was determined if replication in HeLa cells led to the selection of adaptive mutations that were not discovered previously when cDNA clones from SL1
 25 and SL2 cells were sequenced. For this purpose, cDNA clones were isolated from total RNA obtained with pZS2- and pZS25-derived cell lines, respectively. Nucleotide sequence analysis of both cDNA clones did not reveal any additional consensus mutations, suggesting that the two
 30 subgenomes were sufficiently adapted for replication in HeLa cells (results not shown). However, as mentioned above, it was possible that a minor population of subgenomic replicons with additional mutations were present in these cell lines. To overcome this problem, a
 35 method for the isolation and cloning of cDNAs spanning the NS3 to NS5B region was developed (see Materials and Methods). Replicon cDNA libraries were produced from GS4.1, SL1, and MH4 cells. Approximately 2,000 cDNA

clones were pooled and subsequently used for in vitro transcription of subgenomic RNA. With Huh7 cells, the colony formation efficiency of the pooled clones was comparable to that of the most efficient subgenomes, such 5 as pZS2 or pZS25, and did not vary significantly with the origin of the total RNA used for cDNA cloning (Table 4). Consistent with previous results, colony formation in HeLa and mouse cells was origin dependent, i.e., save for one case, colonies were observed only with clones derived 10 from SL1 and MH4 cell lines. Notably, with this strategy G418-resistant colonies were obtained for the first time with Hepal-6 cells by using in vitro-transcribed RNA. To confirm the presence of viral RNA, 11 colonies were expanded and Northern blot analysis was performed with 15 total RNA. All 11 RNA samples analyzed contained viral RNA ranging from approximately 0.1 to 1 ng/5 µg of total RNA (results not shown).

Taken together, the results supported the hypothesis that mutations identified in subgenomic replicons 20 expressed in HeLa and mouse cells play a role in adaptation of the replicons to certain cell-type-specific conditions.

DISCUSSION

25 HCV is known as a species- and tissue-specific virus. The results described herein show that replication of HCV can occur in cells derived from tissues other than liver, indicating that cellular factors required for RNA replication are expressed in cell types other than 30 hepatocytes. One interpretation of this result is that the apparent tropism of HCV for hepatocytes is determined primarily at the level of virus entry or assembly or, alternatively, that HCV can infect many other tissues but has escaped detection due to very low amounts of RNA 35 replication or accumulation. Extrahepatic tissues could

serve as reservoirs for HCV that, as with human immunodeficiency virus, could provide a source of viruses that are refractory to antiviral therapy and, importantly, can be responsible for infection of liver grafts following orthotopic liver transplantation (Chun, T. W., et al., 2002. *J. Infect. Dis.* 185:1672-1676; Laskus, T., et al., 2002. *J. Infect. Dis.* 185:417-421). Such a scenario would have profound implications for antiviral therapy. For example, the targeting of drugs to secondary sites of viral replication and the analysis of drug metabolism in cells other than hepatocytes would become important factors for the development of successful antiviral therapies.

It is conceivable that HCV quasispecies in hepatocytes and other tissues exhibit differences in their composition due to the selection of variants with cell-type-specific adaptations. As shown in this Example, replication of subgenomes in HeLa cells led to the accumulation of clusters of mutations in the NS3, NS4B, and NS5A regions including a deletion in NS5A (Figure 2). Mutations and deletions in NS5A have been found previously in genomes that replicated in Huh7 cells, which could suggest that expression of the natural form of this protein in cell culture somehow interferes with RNA replication (Blight, K. J., et al. 2000. *Science* 290:1972-1975; Guo, J. T., et al., 2001., *J. Virol.* 75:8516-8523; Ikeda, M., et al., 2002. *J. Virol.* 76:2997-3006; Lohmann, V., et al., 2003. *J. Virol.* 77:3007-3019; Lohmann, V., et al., 2001. *J. Virol.* 75:1437-1449). However, mutations in the amino terminus of NS4B have previously not been observed. Notably, in both SL1 and SL2 cells, the mutations changed two or one glutamine residues, respectively, to one of the two basic amino acids arginine and histidine. Moreover, the mutation V1749A was present in all five cell lines examined (Table

2 and Figure 2). Thus far, these results show that these mutations appear to be required for replication in HeLa cells, because only replicons pZS2 and pZS25 carrying these mutations yielded colonies after transfection with
5 in vitro-transcribed RNA (Tables 3 and 4). The amino terminus of NS4B is predicted to reside on the cytoplasmic side of endoplasmic reticulum membranes and may interact with other host or viral proteins required for RNA replication (Hugle, T., et al., 2001. Virology 10 284:70-81). As an integral endoplasmic reticulum membrane protein, NS4B might provide a scaffold for the assembly of replication complexes and act as a regulator for RNA replication. More importantly, a recent study revealed that NS4B can induce particular membrane structures,
15 called membranous webs, proposed to be the site for HCV replication (Egger, D., et al., 2002. J. Virol. 76:5974-5984). Interestingly, genetic analyses with an HCV-related pestivirus identified the amino-terminal region of NS4B as a determinant for cytotoxicity caused by high
20 levels of virus replication (Qu, L., et al., 2001. J. Virol. 75:10651-10662).

In summary, this example demonstrates that HCV RNA replication is not restricted to the human hepatoma cell line Huh7 but instead occurs in HeLa cells and hepatoma
25 cells derived from mice. These findings further facilitate development of a mouse model for HCV infection.

Example II RNA Polymerase Inhibitors Inhibit HCV
30 replication in Transformed HeLa Cells

In this example, the anti-HCV activity of 2'-C-methyladenosine (2CMA, an HCV RNA polymerase inhibitor) was tested on GS4.1 (Huh7) cells, and SL1 (HeLa). The
35 cells were treated with 10 μ M 2CMA. The cells were

harvested at 6, 12, 24, 48, and 72 hours. Total cellular RNA was extracted and viral RNA (vRNA) analyzed by Northern blot analysis. These results indicate that 2CMA effectively inhibits HCV in GS4.1 and SL1 cells (Figure 5 5).

Next, the antiviral activity of the HCV RNA polymerase inhibitor 5-OH-cytidine was tested. GS4.1 (HuH7) cells and SL1 (HeLa cells were treated with the indicated amounts of 5-OH-cytidine. The DNA polymerase 10 inhibitor 5-OH-deoxy-cytidine was used as a negative control. The cells were harvested 72 hours after incubation with the drugs. Total cellular RNA was extracted and viral RNA analyzed by Northern blot analysis. The intensity of the bands corresponding to HCV 15 RNA was determined with a Fuji phosphoimager. These results indicate that 5-OH-cytidine effectively inhibits HCV replication in GS4.1 and SL1 cells (Figure 6).

Example III Cytopathic and NonCytoPathic Responses
20 in Cells Expressing Hepatitis C Virus

Currently, combination treatment with alpha-interferon and ribavirin is the therapy of choice for HCV infection. But this treatment is not always effective, and other treatment choices are limited, or have unproven 25 efficacy. Study of the mechanism of action of IFN- α may help elucidate a new, effective treatment for HCV, or help determine what makes HCV treatment effective.

DNA microarray studies revealed that the antiviral response induced by IFNs alters the expression of 30 hundreds of genes and, hence, is far more complex than previously anticipated (Der, S. D., et al., 1998. Proc. Natl. Acad. Sci. USA 95:15623-15628). Little is known about the nature of the cellular proteins that specifically target viral components and, hence, are 35 responsible for the inhibition of viral replication in

the absence of cell death. In contrast, the major signal transduction pathways required for the innate immune response against many viruses have been elucidated. The first wave of IFN-induced genes depends on the 5 phosphorylation of STAT1 and STAT2 and their interaction with IRF9 (p48) to form the transcription factor complex ISGF3. In addition, viral double stranded RNA (dsRNA) and other unknown viral factors are believed to play an important role in the establishment of an antiviral state. They can activate dsRNA-dependent enzymes such as 10 protein kinase R (PKR) and 2',5'-oligoadenylate synthase (OAS), as well as other still-elusive protein kinases (Smith, E. J., et al., 2001. *J. Biol. Chem.* 276:8951-8957). IFN- α can induce a noncytopathic antiviral 15 response or, alternatively, trigger apoptotic programs leading to the elimination of infected cells (Tanaka, N., et al., 1998. *Genes Cells* 3:29-37).

Nucleotide sequence analyses of HCV genomes isolated from Japanese patients revealed a correlation between the 20 presence of mutations in a short segment of NS5A, termed the IFN sensitivity-determining region (ISDR), and resistance to antiviral therapy with IFN- α (Enomoto, N., et al., 1995. *J. Clin. Investig.* 96:224-230; Enomoto, N., et al., 1996. *N. Engl. J. Med.* 334:77-81). Subsequently, 25 it was reported that the ISDR motif can bind to PKR (Gale, M. J., Jr., et al., 1997. *Virology* 230:217-227). Importantly, the ISDR from IFN-resistant, but not from 30 IFN-sensitive, HCV isolates appeared to be a substrate for PKR, suggesting that IFN treatment of chronically infected patients can lead to the selection of HCV variants with ISDRs that can bind and inactivate PKR (Gale, M. J., Jr., et al., 1998. *Clin. Diagn. Virol.* 10:157-162; Tan, S. L., and M. G. Katze. 2001. *Virology* 284:1-12).

35 Accordingly, study of HCV variants and the pathway

by which IFN inhibits HCV is necessary to provide new HCV treatments, and to prevent selection of IFN- α resistant variants.

5 **MATERIALS AND METHODS**

Chemicals and reagents. Recombinant IFN- α 2b (intron A) was purchased from Schering-Plough. Cycloheximide, 2-aminopurine (2-AP), genistein, sodium salicylate, and wortmannin were obtained from Sigma. SB 203580, PD 98059, 10 vanadate, PP2, rapamycin, and lactacystin were obtained from Calbiochem. Epoxomicin was obtained from Boston Biochem, and caspase inhibitor ZVAD- fluoromethyl ketone (ZVAD-FMK) was obtained from Enzyme Systems Products.

15 Cell culture. Huh7 and HeLa cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, penicillin G, streptomycin, nonessential amino acids, and L-glutamine. For the cell 20 lines carrying HCV and Kunjin virus replicons, 500 μ g of G418/ml was added to the medium. The GS4.1 cell line was derived from a subclone of FCA1 cells as described previously (Guo, J. T., et al., 2001. J. Virol. 75:8516-8523). SL1 is a HeLa cell line expressing HCV subgenomic 25 replicon I_{3'7'NS3-3'} (Lohmann, V., et al., 1999. Science 285:110-113; Zhu, Q., et al., 2003. J. Virol. 77:9204-9210). The KUNCD20 cells represent a pool of approximately 200 colonies of G418-resistant HeLa cells obtained after transfection with the Kunjin virus replicon C20DXrepNeo RNA (Khromykh, A. A., and E. G. Westaway. 1997. J. Virol. 71:1497-1505) (kindly provided by A. Khromykh, Sir Albert Sakzewski Virus Research Center, Brisbane, Australia).

30 Plasmids. pCMV-E3L expressing the vaccinia virus E3L protein was obtained from Robert Schneider, New York

University, New York. pln035 expressing virus-associated (VA) RNA was obtained from David Lazinski, Tufts University, Boston, Mass. pEF-HA-HPIV2 expressing the V protein of human parainfluenzavirus 2 (HPIV2) was obtained from Curt Horvath, Mount Sinai School of Medicine, New York, N.Y. To obtain cDNA clones of the gene encoding human Mx-1, Huh7 cells were treated with 100 IU of IFN- α /ml for 6 h and total cellular RNA was extracted with TRIzol reagent (Invitrogen) and first-strand cDNA was made with an oligo(dT)₁₂₋₁₈ primer and Superscript II DNA polymerase (Invitrogen) by following the manufacturer's direction. For the amplification of Mx-A cDNA, the primers used were 5'-AGTATCGTGGTAGAGAGCTGC-3' (SEQ ID NO:15) and 5'-TAATACGACTCACTATAGGGATGTGGCTGGAGATGC-3' (SEQ ID NO:16). The purified PCR fragment was cloned into the pGEM-T Easy vector (Promega). The identity of the cloned fragment was verified by nucleotide sequence analysis.

RNA extraction and Northern blot hybridization. Total cellular RNA was extracted with TRIzol reagent (Invitrogen) by following the manufacturer's direction. Five micrograms of total RNA was fractionated on a 1% agarose gel containing 2.2 M formaldehyde and transferred onto nylon membranes. Membranes were hybridized with riboprobes specific for plus-stranded HCV replicon RNA and Mx-A and β -actin mRNA in the conditions described previously (Guo, J. T., et al., 2001. J. Virol. 75:8516-8523).

Detection of eIF-2 α phosphorylation by Western blotting. For Western blot analysis of eIF-2 α phosphorylation, cells were treated with 100 IU of IFN- α /ml for 12 h and then transfected with 2 μ g of poly(I:C) per 60-mm-diameter plate by using Lipofectamine (Invitrogen). After

3 h of incubation, cells were lysed in high-salt radioimmunoprecipitation assay buffer (50 mM Tris-HCl [pH 8.0], 250 mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% sodium dodecyl sulfate [SDS]). Proteins (40 µg) were
5 separated on SDS-10% polyacrylamide gel and electrophoretically transferred to nitrocellulose membranes. Membranes were incubated with 50% methanol, washed extensively with water, and blocked with 3% casein in TNET buffer (10 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1 mM EDTA, 0.05% Tween 20). Membranes were incubated with rabbit polyclonal antibodies against eIF-2 α (a gift from Robert Schneider, New York University) or phosphorylated eIF-2 α (eIF-2 α -P; Research Genetics, Inc.) diluted in blocking solution for 1 h and then washed extensively
10 with TNET buffer. Membranes were then incubated with horseradish peroxidase-conjugated goat anti-rabbit and immunoglobulin G (IgG) (Amersham), respectively. The bound IgG was detected with Super-Signal
15 chemiluminescence reagents (Pierce).

20 RPA. For the analysis of IFN- β gene expression, cells were treated with 100 IU of IFN- α /ml for 12 h and then transfected with 2 µg of poly(I:C) per 60-mm plate by using Lipofectamine (Invitrogen). After 3 h of
25 incubation, total cellular RNA was extracted with TRIzol reagent (Invitrogen), and IFN- β mRNA levels were determined by RNase protection assay (RPA) with the help of the RPAII kit purchased from Ambion. The probes complementary to IFN- β (GenBank accession no. M25460) and
30 β -actin (GenBank accession no. BC013380) mRNAs spanned positions 272 to 650 and 1030 to 1250, respectively.

Annexin V-FITC staining. SL1 cells were plated on coverslips in six-well plates 16 h prior to treatment.
35 Cells were then mock treated or treated with 100 IU of

- IFN- α /ml in the absence or presence of 20 μ M ZVAD-FMK. Coverslips were then put on slides and incubated with 100 μ l of staining solution containing annexin V-fluorescein isothiocyanate (FITC) at room temperature for 10 to 15 min. After extensive washes with phosphate-buffered saline, slides were examined with a Nikon fluorescence microscope and photographed with a charge coupled device camera.
- Flow cytometry analysis. To determine the fraction of apoptotic cells, the annexin V assay system (Roche Diagnostics GmbH) was used. Cells were incubated with IFN- α (100 IU/ml) alone or together with 20 μ M ZVAD-FMK for 24 h. The culture medium containing detached cells was collected, and the adherent cells were trypsinized and then combined with the detached cells. The cells were collected by centrifugation and washed once in phosphate-buffered saline. Pelleted cells were resuspended in binding buffer and were incubated with annexin V-FITC at room temperature for 10 to 15 min. The stained cells were then diluted with binding buffer and analyzed by flow cytometry (FACScan; Becton Dickinson).

RESULTS

- IFN- α can induce noncytopathic and cytopathic antiviral responses in cells comprising HCV replicons.

As set forth above, stable HeLa cell lines were established that express HCV subgenomes with an efficiency similar to that of Huh7 cells (Zhu, Q., et al., 2003. J. Virol. 77:9204-9210). Examination of the IFN- α response in HeLa derived cell lines such as SL1 revealed a very similar dose dependent reduction of virus replication. The IC₅₀ of IFN- α in HeLa cell lines was generally in the range of 0.1 IU/ml, approximately 10-fold lower than the IFN- α IC₅₀ in Huh7-derived cell lines

(Figure 7). However, in marked contrast to observations made with Huh7 cell lines, treatment of SL1 and other HeLa derived cell lines with more than 30 IU of IFN- α /ml induced cell death in a significant fraction of cells
5 between 6 and 20 h post treatment (Figure 8A). Cell death was caused by apoptosis, as determined by annexin V staining, and could be prevented by the caspase inhibitor ZVAD-FMK. The fraction of apoptotic cells was determined before and after treatment with the cytokine. The results
10 showed that IFN- α induced apoptosis in more than 30% of SL1 cells compared with 6 to 7% in untreated SL1 and parental HeLa cells (Figure 8B). Several other HeLa-derived cell lines were examined to assure that the results obtained with SL1 cells reflected a general
15 property of HeLa cells expressing HCV subgenomes. Moreover, to test more directly whether IFN-induced apoptosis was caused by viral replication, two methods were used to inhibit RNA replication in SL1 cells. The first was based on the observation that replication of
20 HCV subgenomes is temperature sensitive and is inhibited at 39°C (J. A. Sohn and C. Seeger, unpublished observations). The second method relied on the availability of an inhibitor of the viral RNA polymerase. Consistent with a role for viral replication in the
25 induction of apoptosis, cell death could be prevented when viral replication was inhibited by either incubation of the cells for 60 h at the elevated temperature or treatment with a viral polymerase inhibitor (Figure 8B and result not shown).

30 These results raised the question of whether IFN-induced apoptosis reflected a general property of HeLa cells expressing viral replicons. To address this problem, the IFN response against HCV in SL1 cells was compared with that against the flavivirus Kunjin virus in
35 HeLa cells. For this purpose a pool of HeLa cells,

KUNCD20, expressing Kunjin virus subgenomic replicons lacking the structural genes, similar to the HCV subgenomic replicons was established (Khromykh, A. A., and E. G. Westaway. 1997. J. Virol. 71:1497-1505). The 5 Kunjin virus RNA levels in these cells were approximately fivefold higher than those observed with HCV in SL1 cells. In contrast to results with SL1 cells, treatment of KUNCD20 cells with different concentrations of IFN- α only slightly inhibited viral replication (Figure 9).
10 Importantly, cell death in IFN- α - treated KUNCD20 cells was not detected either by light microscopy or annexin V staining (results not shown). These results indicated that IFN-induced apoptosis is a property of HCV-expressing HeLa cells rather than a general property of
15 HeLa cells replicating viral RNA genomes.

In summary, these results demonstrated that IFN- α could induce noncytopathic as well as cytopathic antiviral programs in cells expressing HCV replicons in a concentration- and cell type-dependent fashion. Moreover, 20 the results showed that this antiviral program was specific for HCV replicons. Importantly, the results suggested that HCV replication could induce an innate cellular response that, in combination with IFN- α , could lead to apoptosis.
25

IFN- α inhibits HCV replication through the Jak-STAT signal transduction pathway

Information about the signal transduction pathways responsible for execution of the IFN response has 30 generally been obtained with cells treated with high concentrations (100 to 1,000 IU/ml) of the cytokine and with fibroblasts and epithelial cells, most of which cannot, to date, support HCV replication. Moreover, a recent study by Schlaak, et al. (Schlaak, J. F., et al., 35 2002., J. Biol. Chem. 277:49428-49437) revealed that the

IFN response could vary in a cell type-dependent manner. In addition, it was found that slight changes in cell culture conditions had major effects on HCV replication. Therefore, the observation that replication of HCV in both Huh7 and HeLa cells could be inhibited with low concentrations of the cytokine warranted a more careful study of the pathways involved in the antiviral program against HCV.

To investigate the nature of the IFN- α response against the HCV replicon, drugs that were known to inhibit specific components of selected signal transduction pathways that play a role in the antiviral response induced by IFN- α were used (Table 5). The current model for the signal transduction pathway induced by IFN- α predicts that the IFN receptor associated tyrosine kinases Jak1 and Tyk2 are activated and, in turn, phosphorylate the transcription factors STAT1 and STAT2, which are required for the induction of the cellular antiviral program (Figure 10) (Sen, G. C. 2001., Annu. Rev. Microbiol. 55:255-281; Stark, G. R., et al., 1998. Annu. Rev. Biochem. 67:227-264). Incubation of GS4.1 cells with the tyrosine kinase inhibitor genistein suppressed the induction of the IFN-induced Mx-1 gene (Figure 11A and 11B). Similarly, genistein antagonized the IFN- α response against the HCV replicon. An increase in the concentration of the drug from 100 to 300 μ M led to a complete inhibition of the IFN response against HCV (results not shown). Consistent with this result, it was found that the V protein of HPIV2 blocked the IFN response. The V protein of HPIV2 induces the degradation of STAT2 and, hence, inhibits the IFN-induced activation of gene expression (Parisien, J. P., et al., 2002. J. Virol. 76:4190-4198) (Figures 11C and 11D). IFN- α treatment of cells expressing HPIV2 led to a twofold reduction of viral RNA levels. When adjusted for the

observed transfection efficiency, i.e., 40 to 45% of the cells express the V protein, the reduction corresponded to a complete inhibition of the IFN- α response. Finally, the decline of viral RNA levels was reduced in GS4.1
5 cells with IFN- α and cycloheximide, indicating that de novo protein synthesis was required for an antiviral response against HCV replication (Table 5).

IFN- α can activate, in addition to the STAT pathway, MAPKs, including extracellular signal-regulated kinase,
10 p38 MAPK, and phosphatidylinositol 3 (PI3)-kinase-Akt pathways (David, M., et al., 1995. Science 269:1721-1723; Goh, K. C., et al., 1999. EMBO J. 18:5601-5608; Pfeffer, L. M., et al., 1997. Science 276:1418-1420). However, in contrast to genistein, SB 203580, sodium salicylate, and
15 wortmannin, known inhibitors of p38 MAPK, NF- κ B, and PI-3 kinase, respectively, did not inhibit the IFN response at detectable levels, suggesting that the two major ancillary signaling pathways activated by IFN- α were not directly involved in inhibiting HCV replication in Huh7
20 cells (Table 5; results not shown).

In summary, the results showed that inhibition of HCV replication with IFN- α depended on a functional Jak-STAT pathway (Figure 10). Hence, the results demonstrated that the IFN response against HCV was genuine and did not
25 reflect an unspecific effect of the cytokine.

Does HCV replication induce an antiviral state in infected cells? A critical step in activation of innate immunity is the induction of an antiviral state by dsRNA or viral proteins (Figure 10) (Taniguchi, T., and A. Takaoka. 2002. Curr. Opin. Immunol. 14:111-116; tenOever, B. R., et al., 2002. J. Virol. 76:3659-3669). As shown above, evidence for such a virus-induced activation was also obtained from IFN-treated HeLa cells expressing HCV
35 replicons. To investigate the nature of this HCV-induced

activation, the phosphorylation levels of eIF-2 α and expression of IFN- β were determined. eIF-2 α is a substrate of PKR, which is activated by dsRNA that can accumulate as a consequence of viral RNA replication

5 (Srivastava, S. P., et al., 1998. J. Biol. Chem. 273:2416-2423; Williams, B. R. 3 July 2001, posting date. Signal integration via PKR. Sci STKE 2001:RE2. [Online.]). IFN- β gene transcription is activated through the coordinate actions of three families of transcription

10 factors NF- κ B, IRF3, and ATF2, all of which are activated by dsRNA and/or certain viral proteins (Figure 10) (Peters, K. L., et al., 2002. Proc. Natl. Acad. Sci. USA 99:6322-6327; tenOever, B. R., et al., 2002. J. Virol. 76:3659-3669).

15 First the levels of eIF-2 α -P in Huh7, GS4.1, HeLa, and SL1 cells was determined in the presence and absence of dsRNA and IFN- α . The results showed that eIF-2 α -P levels were not significantly elevated in cells expressing HCV replicons (GS4.1 and SL1) compared with those in their parental cells (Huh7 and HeLa) (Figure 12A). Similarly, incubation of cells with IFN- α did not induce eIF-2 α -P levels. In contrast, transfection of cells with poly(I:C), mimicking dsRNA, augmented eIF-2 α -P levels, particularly in HeLa and SL1 cells. Similar

20 25 results were obtained when cells were primed with IFN- α prior to transfection with poly(I:C). These results were confirmed with several other cell lines derived from Huh7 and HeLa cells.

Second, the levels of IFN- β mRNA in the four cell lines was determined under the same conditions described above. In agreement with the above results, viral replication alone was not sufficient to activate IFN- β gene expression in both Huh7- and HeLa-derived cell lines (Figure 12B). In Huh7 cells and GS4.1 and other Huh7-derived cells expressing HCV replicons, only a weak

induction of IFN- β was observed when cells were primed with IFN- α and then transfected with poly(I:C). In contrast, IFN- β transcription was induced in HeLa and SL1 cells by poly(I:C) alone and particularly in combination with IFN- α . Remarkably, expression of IFN- β could be induced by IFN- α alone in SL1 cells but not in HeLa cells. Similar results were obtained with the HeLa-derived SL2 cell line (results not shown).

In summary, the results showed that, while both Huh7 and HeLa cells were competent to activate dsRNA-dependent signal transduction pathways, HCV replication alone was not sufficient to induce a detectable dsRNA response in these cells. This result could indicate that dsRNA either does not accumulate during HCV replication or cannot access PKR and other dsRNA binding proteins. Importantly, the results showed that, despite the apparent lack of biologically active dsRNA, viral replication could activate certain cellular signal transduction pathways that could cooperate with IFN- α to activate the transcription of the IFN- β gene.

The results presented above favored a model predicting that IFN- α inhibited HCV replication by a mechanism that was independent of dsRNA-activated antiviral pathways. To test this model more carefully, the IFN response was measured in GS4.1 cells expressing the vaccinia virus E3L protein. E3L is known to sequester dsRNA and prevent PKR and OAS/Rnase L activation (Chang, H. W., et al., 1992. Proc. Natl. Acad. Sci. USA 89:4825-4829; Rivas, C., et al., 1998. Virology 243:406-414) (Figure 10). Indeed, expression of E3L had no measurable effect on the IFN response against HCV (Figure 11C and 11D). Experiments relying on simultaneous detection of E3L and the HCV protein NS5A in the same cell by immunofluorescence confirmed that IFN- α could inhibit HCV replication in cells expressing the E3L protein (results

not shown). Finally, it was found that the PKR inhibitors 2-amino purine (2-AP) and adenovirus VA RNA did not block the IFN response against HCV in Huh7 cells (Table 5).

5 TABLE 5. Effects of inhibitors on the activity of IFN- α
against the HCV replicon^a

	Drug, protein, or RNA	Concn	Primary target(s)	Effect
10	2-AP	10 nM	PKR and other kinases	-
	Genistein	300 μ M	Tyrosine kinases	+
	Cycloheximide	10 μ g/ml	Translation	+
	Sodium salicylate	5 mM	IKK	-
15	SB 203580	20 μ M p38	MAPK	-
	PD 98059	50 μ M	MEK kinase	-
	Vanadate	50 μ M	Protein phosphatase	-
	Wortmannin	100 nM	PI3 kinase	-
	PP2	50 μ M	src kinase	-
20	Rapamycin	200 nM	mTOR, translation	-
	Lactacystin	5 μ M 26S	proteasome	+
	Epoxyomicin	1 μ M 26S	proteasome	+
	V protein of HPIV2		STAT2	+
	E3L protein		PKR and OAS	-
25	VA RNA		PKR	-

* GS4.1 cells were incubated with the indicated compounds for 2 h and then the presence of 100 IU of IFN- α /ml for an additional 24 h. Viral RNA levels were determined by Northern blot analysis. Assays for V protein, E3L, and VA RNA are described in the legend to Fig. 5.

What are the pathways that play a role in the IFN- α response against HCV? A major question concerns the mechanism by which IFN- α induces the noncytopathic inhibition of HCV replication. DNA microarray analyses of IFN- α -treated GS4.1 cells and other Huh7-derived cell lines revealed the induction of several classes of genes belonging to known signal transduction and protein degradation pathways (J. Hayashi and C. Seeger, 35 unpublished results; Cheney, I. W., et al. 2002. J. Virol. 76:11148-11154). In particular, several genes encoding proteasome subunits and ubiquitin-like proteins were among the genes most highly induced by IFN- α .

Notably, kinetic studies of HCV replication in HuH7 cells indicated that replication complexes have a relatively short halflife, which is further reduced by IFN treatment (J.-T. Guo and C. Seeger, unpublished observations).

- 5 Therefore, it was surmised that the proteasome could play a role in inhibition of HCV replication in IFN-treated cells.

To test this hypothesis, the outcome of combination treatment with IFN- α and the proteasome inhibitors 10 lactacystin and epoxomicin for HCV RNA replication in GS4.1 cells was determined. The cells were pretreated with different concentrations of the inhibitors for 1 h before IFN- α was added for an additional 6 h of combination treatment. Then the cells were incubated for 15 12 h before RNA was isolated and subjected to Northern blot analysis (Figure 13A). The relatively short incubation period was necessary because of the known toxicity of proteasome inhibitors after longer incubation times. The results showed that HCV RNA levels dropped 70% 20 within 18 h of IFN- α treatment, whereas in the presence of epoxomicin or lactacystin the reduction was only 30% (Figure 13B). Lower doses of epoxomicin than of lactacystin were effective, which is consistent with the high specific activity of epoxomicin against the 25 chymotrypsin-like activity of proteasomes (Fenteany, G., and S. L. Schreiber. 1998. J. Biol. Chem. 273:8545-8548; Meng, L., et al., 1999. Proc. Natl. Acad. Sci. USA 96:10403-10408). Treatment with higher doses of lactacystin alone led to a slight reduction of HCV RNA 30 levels. These results were confirmed with a second set of experiments. The cells were pretreated with 5 μ M lactacystin and 1 μ M epoxomicin, respectively, for 1 h before IFN- α was added for an additional 6 h of combination treatment. RNA was isolated from the treated 35 cells either 6 or 12 h after incubation with IFN- α .

(Figure 14). The results showed that, at both time points, viral RNA levels were significantly higher in cells that were exposed to the proteasome inhibitors than in cells that were treated with IFN- α alone.

5 Finally, tests were conducted to determine whether proteasome activity was required for induction of the IFN response or, more directly, for inhibition of HCV replication. To distinguish between these two possibilities, first GS4.1 cells were incubated with the
10 cytokine for 10 h to induce the antiviral response. Then, the cells were incubated for 12 h in the presence of lactacystin or epoxomicin. Under these conditions, the IFN response against HCV remained effective and reduced RNA levels to similar extents independently of the
15 presence of either of the two inhibitors (Figure 15). Thus, these results indicated that the activity of proteasomes is required for the induction of the IFN response against HCV, but apparently not for direct inhibition of viral replication (see Discussion).

20

DISCUSSION

In this Example, the mechanism of the IFN- α response against subgenomic replicons of HCV in Huh7 and HeLa cells is investigated. The following conclusions can be
25 drawn from these investigations. First, it can be concluded that IFN- α can inhibit HCV replication by both noncytopathic and cytopathic mechanisms. These results demonstrating that SL1 cells treated with IFN- α (100 IU/ml) underwent programmed cell death raised the
30 question of whether apoptosis contributes to the rapid decline of HCV RNA levels observed during the first 48 h of IFN therapy (Neumann, A. U., et al., 1998. Science 282:103-107). The answer depends on whether HeLa or Huh7 cells mimic the scenario in HCV-infected hepatocytes in
35 vivo. It is known from this and other studies that Huh7

cells exhibit an attenuated response to dsRNA and cannot induce an apoptotic program (results not shown) (Keskinen, P., et al., 1999. *Virology* 263:364-375; Lanford, R. E., et al., 2003. *J. Virol.* 77:1092-1104; 5 McNair, A. N., et al., 1994. *J. Gen. Virol.* 75:1371-1378). In contrast, HeLa cells respond to dsRNA in a fashion similar to that in which primary chimpanzee hepatocytes respond. For example, treatment of chimpanzee primary hepatocyte cultures with poly(I:C) led to the 10 induction of IFN- β , as shown in this report with HeLa cells (Figure 12) (Lanford, R. E., et al., 2003. *J. Virol.* 77:1092-1104). Therefore, it is likely that HeLa cells represent a more physiological model for hepatocytes in terms of IFN response than Huh7 cells. It 15 was notable that only a fraction of SL1 cells died after treatment with IFN- α . One possibility is that apoptosis was induced in cells that replicated above-average levels of HCV RNA. In support of this possibility, reduction of viral levels by treatment with heat or HCV RNA polymerase 20 inhibitors reduced the number of apoptotic cells after IFN- α treatment (Figure 8 and results not shown). Based on these results it was concluded that HeLa cells did not undergo apoptosis by default after IFN- α treatment. In fact, it appears that apoptosis is a hallmark of HCV- 25 replicating HeLa cells, because HeLa cells replicating Kunjin virus RNA remained viable after IFN treatment. Finally, it appears that, the observation reported here represents the first example of IFN- α -induced apoptosis of cells replicating an apparently noncytolytic RNA 30 virus.

Second, it can be concluded that the noncytopathic response can occur independently of dsRNA-dependent pathways. Although these results showed that dsRNA response pathways were at least partially functional in 35 normal and HCV-replicating Huh7 cells and were intact in

HeLa cells, as indicated by poly(I:C)-induced phosphorylation of eIF-2 α and IFN- β gene transcription, viral replication per se did not induce such responses (Figure 12). The expression of the vaccinia virus E3L 5 protein or treatment of cells with the kinase inhibitor 2-AP had no measurable effect on the IFN response against the HCV replicon (Figure 11 and Table 5). These observations were consistent with the notion that dsRNA-dependent antiviral pathways, such as PKR and RNase L 10 pathways, were not involved in IFN-induced inhibition of HCV replication in Huh7 cells. Whether they play a role in HeLa cells is not yet known. Efforts to express E3L in SL1 cells were not successful due to the apparent 15 toxicity of the protein, and treatment of HeLa cells with 2-AP for more than 12 h induced apoptosis (results not shown). Importantly, it is not yet known whether IFN-induced apoptosis in HCV-expressing cells is dependent on PKR or other dsRNA-dependent pathways (see below).

In summary, the results showed that, while both Huh7 20 and HeLa cells were competent to activate dsRNA-dependent signal transduction pathways, HCV replication alone was not sufficient to induce a detectable dsRNA response in these cells. This result could indicate that dsRNA either did not accumulate during HCV replication or was not 25 accessible to PKR and other dsRNA binding proteins.

Third, it can be concluded that HCV replication can induce innate immune pathways. In SL1 and other HCV-expressing HeLa cell lines (results not shown), but not 30 in normal HeLa cells, IFN- α induced the expression of IFN- β . This indicates that HCV replication activated an unknown cellular factor, perhaps a viral activated kinase as proposed by Smith and colleagues (Smith, E. J., et al., 2001. J. Biol. Chem. 276:8951-8957), that, in turn, activated one or more transcription factors required for 35 IFN- β transcriptional activation. Candidates include

IRF3, NF- κ B, and ATF-2 (Figure 10). Because normal HeLa cells did not undergo apoptosis after IFN- α treatment, it can be concluded that HCV expression directs the IFN- α response into a cytopathic process.

- 5 Fourth, it can be concluded that the antiviral activity of IFN- α against HCV depends, in part, on functional proteasomes. How IFN-induced antiviral programs inhibit viral replication noncytopathically is not yet understood. The results shown here demonstrate
10 that, for HCV replication, proteasomes were required for this process. However, the idea that proteasomes were directly involved in inhibition of HCV replication by increasing the turnover of replication complexes or viral proteins was not supported by these results. Instead
15 evidence was obtained that induction of the IFN response was dependent on degradation of one or several proteins. Previously, it has been shown (Li, X. L., and B. A. Hassel. 2001. Cytokine 14:247-252) that proteasome inhibitors attenuated the induction of certain IFN-
20 stimulated genes. Because epoxomicin and lactacystin did not inhibit induction of Mx-A (results not shown), which is dependent on activation of the Jak-STAT pathway for the formation of ISGF3, proteasomes might be involved in the induction of the second-wave IFN-stimulated genes.
25 Such a model is consistent with results published previously by Li and Hassel (Li, X. L., and B. A. Hassel. 2001. Cytokine 14:247-252), who found that treatment of cells with proteasome inhibitors did not inhibit phosphorylation of STAT1 and binding of ISGF3 to DNA. As
30 a consequence of these results, the number of IFN-induced genes that play a role in inhibition of HCV replication by IFN- α can be reduced to those that are repressed by epoxomicin.

35 An important implication of these results for clinical IFN- α therapy and the pathogenesis of HCV

infections is that besides the noncytopathic antiviral effects, IFN- α might also induce apoptosis of HCV-infected hepatocytes. At first glance, this possibility might be discounted because drug-induced cell death could 5 lead immediately to the destruction of the infected liver. However, it is possible that only a fraction of hepatocytes express levels of HCV high enough to activate an apoptotic program in the presence of the high levels of IFN- α that are used for antiviral therapy. In this 10 scenario cell death would occur unnoticed. An important consequence of such a scenario would be that cell killing could play a major role in the recovery from chronic HCV infections, similar to the situation in natural recovery from transient infections with woodchuck hepatitis virus, 15 a model for hepatitis B virus infections (Guo, J. T., et al., 2000. J. Virol. 74:1495-1505).

Table 6. Listing of Sequence ID Numbers

Sequence	Sequence ID Number
I ₃₇₇ /NS3-3'	SEQ ID NO:1
pZS1	SEQ ID NO:2
pZS2	SEQ ID NO:3
pZS4	SEQ ID NO:4
pZS5	SEQ ID NO:5
pZS6	SEQ ID NO:6
pZS8	SEQ ID NO:7
pZS10	SEQ ID NO:8
pZS11	SEQ ID NO:9
pZS12	SEQ ID NO:10
pZS15	SEQ ID NO:11
pZS20	SEQ ID NO:12
pZS25	SEQ ID NO:13
pZS45	SEQ ID NO:14
Mx-A cDNA primer #1	SEQ ID NO:15
Mx-A cDNA primer #2	SEQ ID NO:16

While certain preferred embodiments of the present
5 invention have been described and specifically
exemplified above, it is not intended that the invention
be limited to such embodiments. Various modifications
may be made to the invention without departing from the
scope and spirit thereof as set forth in the following
10 claims.

WHAT IS CLAIMED IS:

1. A cell-line which replicates hepatitis C virus (HCV), wherein said cell line is selected from the group 5 consisting of a non-human cell line and a human non-hepatic cell line.
2. The cell line of claim 1, wherein the human non-hepatic cell line comprises epithelial cells.
10
3. The cell line of claim 2, wherein the human epithelial cells are HeLa cells.
4. The cell line of claim 1, wherein the non-human 15 cell line comprises mouse cells of hepatic origin.
5. The cell line of claim 4, wherein the mouse cells are Hepal-6 cells.
- 20 6. The cell line of claim 4, wherein the mouse cells are AML12 cells.
7. A non-human, non-chimpanzee host animal comprising cells which replicate HCV.
25
8. The non-human host animal of claim 7, which is a mouse.
9. A method for producing a human non-hepatic cell 30 that replicates HCV, comprising:
 - a) obtaining total RNA from a human hepatic cell culture that replicates HCV, said total RNA comprising a selection marker which renders cells expressing said RNA resistant to a selection agent;
 - 35 b) introducing the total RNA into human non-

hepatic cells; and

c) selecting those cells which grow in the presence of said selection agent and replicate HCV.

5 10. The method of claim 9, wherein a cell line is generated from the cells of step c).

11. A method of producing a non-human hepatic cell that replicates HCV, comprising:

10 a) obtaining total RNA from a human non-hepatic cell culture that replicates HCV, said total RNA comprising a selection marker which renders cells expressing said RNA resistant to a selection agent;

b) introducing the total RNA into non-human 15 cells; and

c) selecting those cells which grow in the presence of said selection agent and replicate HCV.

12. The method of claim 11, wherein a cell line is 20 generated from the cells of step c).

13. A method for screening test compounds which inhibit HCV replication, comprising:

25 a) culturing the cell line of claim 1 in the presence and absence of a test compound; and

b) assaying HCV replication levels in the presence and absence of said test compound, wherein a reduced HCV replication level in the presence of said test compound is indicative that said test compound 30 inhibits HCV replication.

14. An HCV polynucleotide having at least one of the mutations shown in Table II.

35 15. A polyprotein encoded by the polynucleotide of

claim 14.

16. A method for screening test compounds which modulate the antiviral response induced by interferon 5 alpha (IFN- α) comprising
- a) culturing the cell line of claim 1 in the presence and absence of a test compound;
 - b) contacting the cells of step a) with IFN- α ; and
 - c) measuring the HCV replication level in the 10 presence and absence of said compound thereby identifying agents which modulate the antiviral response mediated by IFN- α as a function of altered HCV levels.

17. The method of claim 16, wherein the antiviral 15 response is enhanced.

18. The method of claim 16, wherein the antiviral response is inhibited.

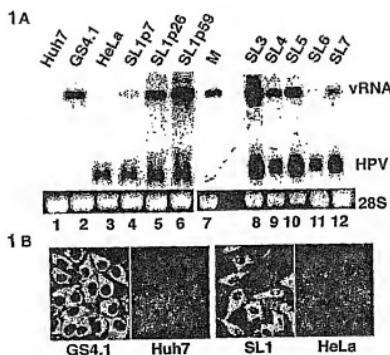
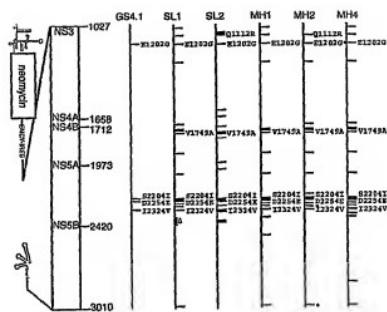
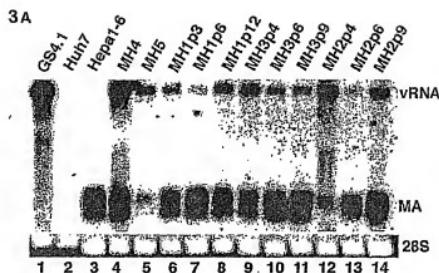
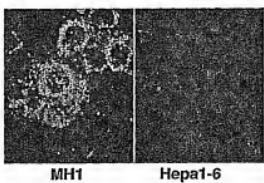
Figures 1A-1B

Figure 2

3/15

Figures 3A-3B

3 B



4/15

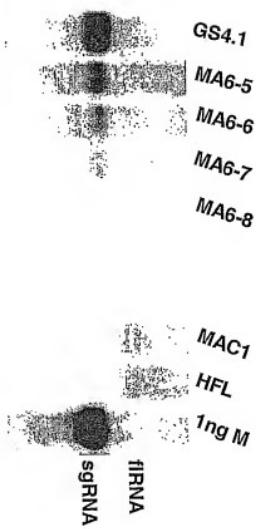


Figure 4

Figures 5A-5B

A

Time (h) 0 6 12 24 48 72

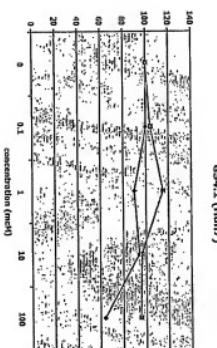
**B**

Time(h) M 0 24 48 72

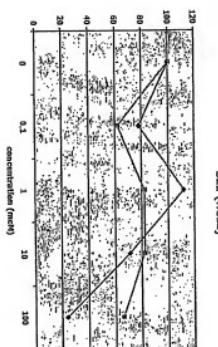


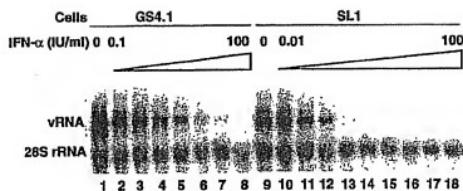
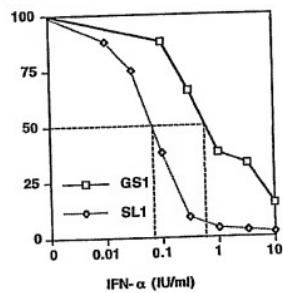
Figures 6A-6B

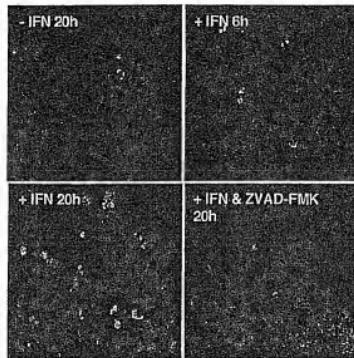
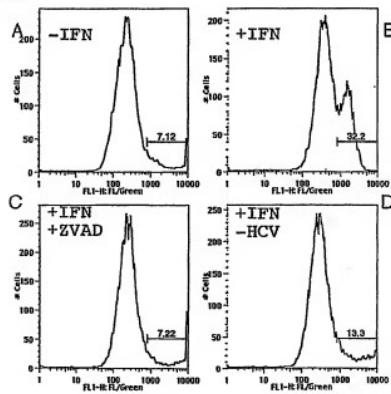
6A



6B



Figures 7A-7B**7A****7B**

Figures 8A-8B**8A****8B**

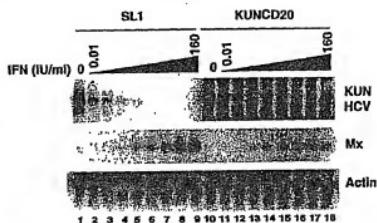
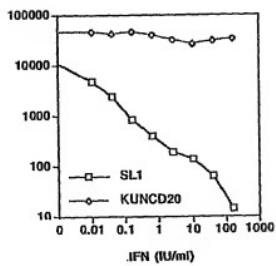
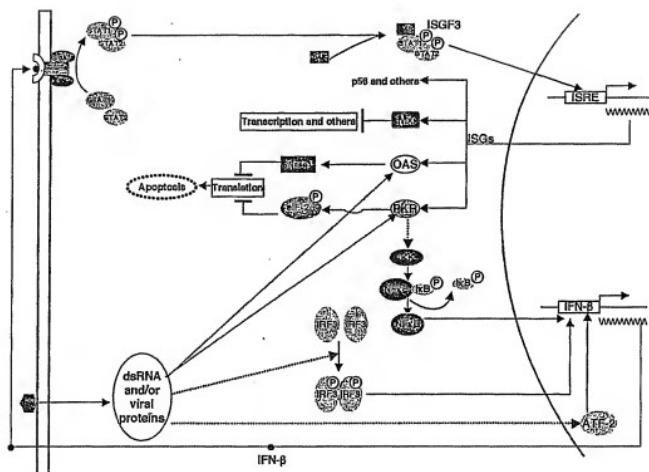
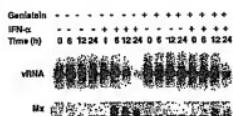
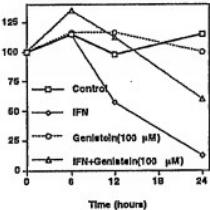
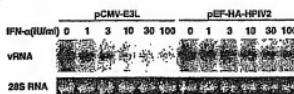
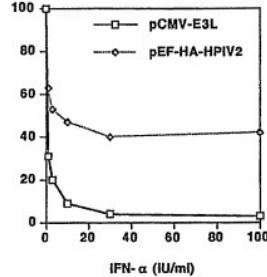
Figures 9A-9B**9A****9B**

Figure 10

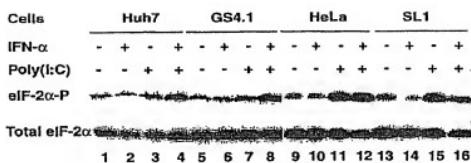
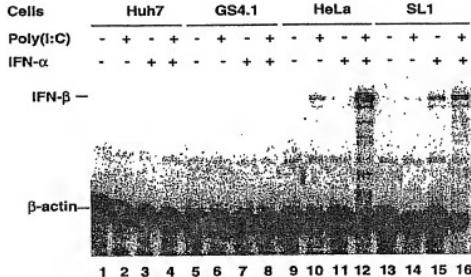
Figures 11A-11D

11A

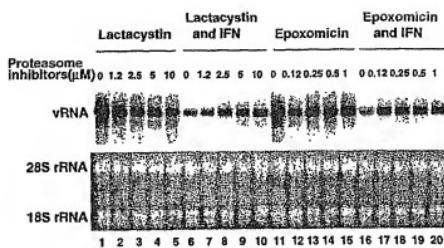
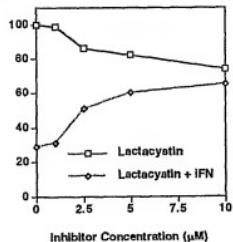
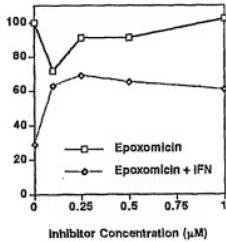
28S RNA [band] [band]

11B**11C****11D**

Figures 12A-12B

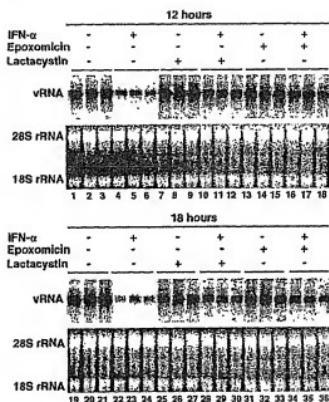
12A**12B**

13/15

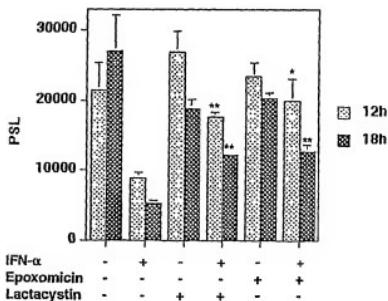
Figures 13A-13C**13A****13B****13C**

Figures 14A-14B

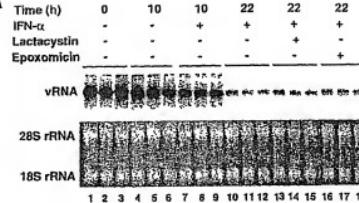
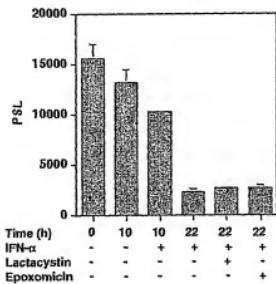
14 A



14B



Figures 15A-15B

15A**15B**

SEQUENCE LISTING

<110> Qing Zhu
Ju-Tao Guo
Christoph Seeger

<120> Replication of Hepatitis C Virus in
Non-Hepatic Epithelial and Mouse Hepatic Cells

<130> 0149 PO3068

<140> Not Yet Assigned
<141> 2003-12-12

<150> 60/433,303
<151> 2002-12-13

<160> 16

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 11313
<212> DNA
<213> Artif

<220>
<223> Plasmid

caggaccctcg	tccggctggca	agcgcgccccc	ggggcgcgtt	ccttgacacc	atgcacccgtc	2100
ggcagctcg	acctttactt	ggtaacggagg	catggccgat	tcattcccggt	gcccggccgg	2160
ggegacaga	gggggagact	actctcccc	aggcccttgc	cctaactgaa	gggttcttcg	2220
ggcggtccac	tgcctcgccc	ctcgggcac	gctgtggca	tcttcgggc	tgccgtgtgc	2280
acccaggagg	tttgcgaagg	ggttggactt	gtacccgtgc	agttctatgg	aaccactatg	2340
cggttccccc	tttcaacgg	caactctgc	ctccggccgc	tacccgcaaa	atcccaatgt	2400
gccccatcac	acggccctac	tggtagccgc	aagacacta	aggtgcgggc	tgcgtatgca	2460
gcccaagggt	ataagggtgt	tgtctgttgc	ccgtccgtcg	ccgcacccct	aggtttcggg	2520
ggtgtatgtt	ctaaggcaca	ttggatccgg	cotaacatca	gaacgggggt	aaggaccatc	2580
accacgggtt	ccccccatca	gtactccacc	tatggcaagt	tttcttcgcga	cggtgttgc	2640
tctggggggc	cctatgcac	ataataatgt	gatgagtgc	actaactgaa	ctcgaccatc	2700
atccctggca	ttccgcacat	cttgcacca	ggggagacgg	cttgcacgcg	actgtcggt	2760
ctcgccaccc	ctacgcgtcc	gggatgggtc	accgtgcac	atcccaatcc	cgaggagggt	2820
gctctgtcca	gcaactgg	aatccccctt	tatggcaaa	ccatccccat	cgagaccatc	2880
aagggggggg	ggccacccat	ttttctcgat	tccaaaga	aatgtgtatg	gctgcgcgcg	2940
aagctgtccg	gcctcggact	caatgtgttgc	gcattttacc	ggggcttgc	tgatctccgc	3000
ataccaacta	ggggagacgt	ttttgtcgta	gcaacggacg	cttcaatgc	gggttttacc	3060
ggcgattttc	actcgtatgc	caactgtca	atagttgtca	ccacagacat	cgacttcacg	3120
ctggggccgc	ctttcacat	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3180
cagggggggg	ggcggactgg	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3240
gaacggccct	cgggcatgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3300
ggttttgttc	agtcgtaccc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3360
ccagggtttcc	ccgtctggca	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3420
acccacatag	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3480
ctggtagatcc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3540
caaaatgttgc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3600
tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3660
atggcatgttcc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3720
gtcttagcag	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3780
atccatctttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3840
gtatggatgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3900
gaacaatccat	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	3960
gtgtccccc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4020
tggaaatttca	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4080
ggttatgtat	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4140
accccttcgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4200
gtttctgttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4260
aagggtttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4320
ttaaagggttcc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4380
atcccttcgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4440
gtggccccc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4500
ggtaaacacg	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4560
cagatctctt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4620
gaggactgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4680
acgggttgg	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4740
gtccctttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4800
tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4860
caaacacact	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4920
atcggtttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	4980
accacggggc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5040
gtctgttgg	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5100
accaactgaca	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5160
gtatgggtttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5220
acatttcgttcc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5280
ccggactgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5340
gtaaagcgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5400
ctgtctgtcc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5460
ctatcgttgg	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5520
tcagaaaaat	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5580
aggaaatgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5640
cccatatgg	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5700
tacgcttcc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5760
cctccacgg	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5820
gagotcgtca	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5880
acggcccttc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	5940
tcctccatgtc	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	6000
ttaacatgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	

acaggcgccc	tgatcacgoc	atgcgctgog	gaggaaacca	agtcgccccat	caatgcactg	6060
agcaactt	tgctcgtca	ccacaatgg	gtctatgtca	caacatctcg	cagcgaacg	6120
ctgcggcaga	agaagggtac	ctttgacaga	ctgcgggtcc	tggacgacca	ctaccgggac	6180
gtgctcaagg	agatgtggc	gaaggcgcc	acgatgttgg	ctaaacttct	atccgtggag	6240
gaaggctgtt	agctgacgoc	ccccatctt	gcccatagtt	aatttgcgtt	tggggcaaa	6300
gacgtccgga	acatcttccag	caaggcggtt	aaccacatcc	gctccgtgtt	gaaggacttg	6360
ctggaaagaca	ctggagacacc	aatttgacacc	accatcatgg	aaaaaatgg	ggttttctgc	6420
gtcacaaccc	agaagggggg	ccgaagggc	gctcgcttta	tgttatttccc	agatttgggg	6480
gttctgtgt	gcgagaaaat	ggccctttac	gatgttgttcc	ccacccctcc	tcaggcgctg	6540
atgggtctt	catacggatt	caatactt	cctggacagc	gggtcgagg	ccttggtaat	6600
gctctggaaat	cgaaagaaatg	ccctatggc	tttgcataat	acaccggct	ttttgtacta	6660
acggtcactt	agataatgtt	ccgtgttg	gagtcataat	accaatgtt	tgactttgg	6720
ccggaaagcca	gagaggccat	aagggtctcc	acagacggc	ttatcatccgg	ggggccccctg	6780
actaaatcta	aaggcgagaa	otggggctat	egccgtgtcc	ggccggccgg	tgactgtcag	6840
accaggctcg	tgatataatcc	catatgtac	ttgaaggccg	otggggctcg	tcagactcg	6900
aaatcttcagg	actgcacatc	gctcgatgt	ggagacggcc	ttgttggat	ctgtgaaagc	6960
ggggggaccc	aaaggaggcga	ggccgtccat	ggggccctcc	cgagggtat	gactatgtac	7020
tctggccccc	ctggggaccc	gccccaaacc	gaatacgtac	ttgggtgtat	aacatcatgc	7080
tcttccaatg	tgtcagtctg	gcacgtatca	tctggccaaa	gggtgtactt	tctcacccgt	7140
gaccggccca	ccccccatgd	gggggttgg	ttggagacac	tttgcacacac	tccagtcata	7200
tcttggtag	gcaacatctt	catgtatgg	ccccatctgt	ggcaaggat	gatctgtat	7260
actcatttct	tctccatctt	ttagatcg	gaacaaatgg	aaaaacccct	agatgttgcag	7320
atctacgggg	ccctgttactt	catgtggcc	cttgcacatc	ctcgcataat	tcacgcactc	7380
catggccctt	ggccatctt	actccatagt	tacttccag	gtgagatcaa	taggttggt	7440
tcatgcctt	ggaaatcttg	gggttggcc	tttgcgttcc	ggagacatcg	ggccggaaatg	7500
gtccgcgtt	ggctactgtt	ccagggggg	agggtggcc	cttggggcc	gtacacttc	7560
aactggcg	taaggacaa	gtccaaatct	actccatca	ccgtgtcg	ccatgttgat	7620
ttatccatgt	ggtttctgt	tggttacagc	ggggggagaca	tatatacag	cctgttctgt	7680
gccccaccc	gggtgttcat	ttgtgtctca	ttccacttct	ctgttaggggt	aggccatctat	7740
ctactccca	acccatgtac	ggggacactt	acacttcagg	ccaaatggcc	atccatgttt	7800
tttccctttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttctctttt	7860
tttttctct	ttttttctt	tttcttctt	ttgtgtctcc	atcttagccc	tagtcacccg	7920
tagtgcgtt	aggtttgcgtt	ggccgttgc	tgccagatgt	gtgtatgtt	gccttttgc	7980
agatcaatgt	ctctgtcagg	ccggccacta	ttggggatcc	ggggggatgt	ocgggtttta	8040
gcataatgt	gacccaaatc	tcatgttgc	cagttatcc	tcgtatgtt	ttatgtgggt	8100
agtttatcac	agtttaatcc	ctaacgcgt	caggcaccc	gtatgtatcc	taacaatgc	8160
ctatcatgtc	tcctccggac	cgtcacccct	gtatgtgtt	gtatgtgtt	gtttagtgcg	8220
gtactgcgg	ggatcttgcg	ggatatactt	catccggaca	gtatgtccgg	tcactatgtc	8280
gtgtgtatgt	cgatatatgt	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	8340
tcggacccgt	ttggccggcc	ccccactctg	ctgggttcc	ttttatgtac	ttttatgtac	8400
tacggcgat	ttggggaccc	ccccgttctg	ttttatgtac	ttttatgtac	ttttatgtac	8460
ggcatcaccc	ggccacccat	ttgggttgc	ttttatgtac	ttttatgtac	ttttatgtac	8520
gaagatcggtt	cttcggccat	ccgggttcat	ttttatgtac	ttttatgtac	ttttatgtac	8580
ggccccgtgg	ccgggggact	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	8640
gggggtgtca	acggccatca	cttactactg	ttttatgtac	ttttatgtac	ttttatgtac	8700
ggagacgtcc	ggccatgtcc	tttggagacc	ttttatgtac	ttttatgtac	ttttatgtac	8760
cgggggatcc	ctatctgtcc	ccggccatct	ttttatgtac	ttttatgtac	ttttatgtac	8820
cagggtccgg	cagggtctcg	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	8880
atgtatccgc	tttgcgttgc	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	8940
actgttccgc	cccaaaaaatc	tttgcgttgc	ttttatgtac	ttttatgtac	ttttatgtac	9000
gacgcgttgc	gtactgtttt	tttgcgttgc	ttttatgtac	ttttatgtac	ttttatgtac	9060
atgatcttc	ttttttttcc	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9120
caggtagatg	acggccatca	ggggacagttt	ttttatgtac	ttttatgtac	ttttatgtac	9180
acttcgatca	ttggggccgt	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9240
aacgggttgc	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9300
cgccgttcat	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9360
gattcaccat	tccaaatgtt	ggggccatca	ttttatgtac	ttttatgtac	ttttatgtac	9420
ccaaaccttg	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9480
tctccggcc	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9540
ccgggttcat	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9600
gaacgttgc	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9660
tccgttccg	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9720
tccggatctt	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9780
cgaaatgttac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9840
ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9900
tgagcatctt	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	ttttatgtac	9960

acggaggcat	caagtgcacca	aacaggaaaaa	aaccgcctt	aacatggccc	gttttatcag	10020
aaggcagaca	ttaacgcttc	tgtagaaact	caacgagctg	gacgoggatc	aacaggcaga	10080
catctgtgaa	tcgcgttca	accacgcgtg	tgagcttac	cgcagctgc	tcgcgcgtt	10140
cggtgtatgg	gggtaaaaacc	tctgcacat	gaatcgcccg	gagacggtca	cagetgtct	10200
gtaaacggat	ggccggggac	gacaacggcc	tcaggccgg	tcaggggttg	ttgggggtt	10260
tcggggggca	gccccatgcac	agtcacgtag	cgatagccg	tgatatactg	gtttaactat	10320
ggggcatggc	agcaggatgt	actggagatg	caccatatgc	ggtgtbaaat	accgcacaga	10380
tgcgtaagg	gaaaataccg	catcaggcg	tttccgtt	cctcgctac	tgactcgctg	10440
cgctcgctg	ttccggatcg	ggagggcgta	tcagctcat	caaaggccgt	aatacgggt	10500
tccacagat	caggggataa	cgaggaaagg	aacatgtgag	caaaaggccg	gcaaaaggcc	10560
aggaaccgt	aaaaggccgc	gtgtctggc	ttttccata	ggctccgccc	ccctgacgag	10620
catcacaaaa	atcgccatca	aagttagagg	ttggcaaaacc	cgacaggact	ataaaagatg	10680
caggcttgc	ccccctggaa	ctccctcg	cgctctctt	ttccgaccc	ggccgttacc	10740
ggatatactg	cecgatcttc	cccttcggg	agcggtggc	ttttctatag	cttcacgtgt	10800
aggtatctca	gttgggtgt	ggctgttgc	tccaaactgt	gctgtgtca	cgaacccccc	10860
gttccagccg	acccgtcg	cttattccgt	aactatgttc	tttagtccaa	cccggttaga	10920
caacgactat	cggccatcg	agcagccact	gttacacagg	tttagcagac	gaggatgtg	10980
ggcggtgtca	caagatctt	ggatgtgtgg	cctaactacg	gctactactg	aaggacagta	11040
tttgcgtatct	ggcgtctgt	gaagccgtt	acccctggaa	aaagatgttg	tagctcttga	11100
ttccggcaaa	aaaccacccg	ttgttaggg	gttttttttt	tttgcgaatg	cgagattacg	11160
cgcgaaaaaa	aaggatctca	agaagatcc	tttgcatttt	ctacgggtc	tgacgctcag	11220
ttggaaacaa	actcacgtt	agggatttt	gtcatgagat	tatcaaaaag	gatcttacc	11280
tagatccccc	tctagataat	acgactact	ata			11313

<210> 2
<211> 11313
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 2						
gccagcccc	gatgggggg	gacactccac	catagatcac	tccctgtga	ggaactactg	60
tcttcacgca	gaaagcgtct	agccatggcg	ttatgtatgg	tgtctgtcag	cctccaggac	120
ccccccctcc	ggggagacca	tagtggctcg	cgaaacccgt	gagtagccac	gaatttgcac	180
gacgacccgg	tcctttctcg	gtacaaacccg	ctcaatggc	ggagatttgg	gctgtcccc	240
gcgagactcg	tagccgatg	gtgtgtggc	ggcaaaaggcc	ttttgtgtat	gcctgtatagg	300
gtgttgcga	gtggccccc	aggctctcg	gacgtgcac	catgagacag	aatcttaaac	360
ctcaaaagaa	aacccaagg	cgccgcattg	ttgaacaaga	ttgatttgc	gcagggtttc	420
cgccgcggc	gggtggaggg	ctatccgt	atgactggc	aaacacagaa	atcggtctgt	480
ctgtatccgc	cggtgttccgg	ctgtgcacgg	agggggccgg	ggttttttt	gtcaagaccc	540
acctgtccgg	tgccctgtat	gaactgtgg	acggaggccg	ggggctatcg	ttggctggcc	600
cgagggggt	tctttgcga	gtgtgtcg	acgtgtgcac	tgaaggggga	agggactgtc	660
tgctatttgc	cgaaatgtcg	gggcaggatc	tccgtcata	ttacccctgt	cctggcgaga	720
aatgtatccat	catggcgat	gcaatggcc	ggctgtatcc	gttgcattcg	gotacctgtcc	780
cattcgatca	ccaaacggaa	catcgatca	agcgacacg	tactcgatcg	gaaggccgtc	840
ttgtcgatca	ggatgtatcg	gacgaagacg	atcagggggt	cgccgcagcc	gaactgttgc	900
ccaggccaa	ggcgcgcatg	cccgaggccg	aggatctgt	cggttgcac	ggcgtatcc	960
gcttcgcgaa	tatcatgtcg	gaaaatggcc	gtttttctgg	attatcgac	tgtggccgg	1020
ttgggtgtgg	ggacgcgtat	caggacatag	cggtggctcg	cgtgtatatt	gctgaaggac	1080
ttggccggcg	atggggctgc	cggttcttcg	ttgtttacgg	tatcgocgt	cccgatccgc	1140
agcgatcg	cttctatcg	tttcttgcg	agtttcttctg	agtttaaaca	gaccacaaacg	1200
gtttccatct	acggcgatcc	attccggccc	tccctccccc	ccccccctaa	cgttactgtc	1260
cgaacgcgt	ttggaaatagg	ccgggtgtcg	ttttgtatca	tgttatttc	caccatattt	1320
ccgtttttgc	gcaatgttg	ggccggggaa	cctggccctg	tttttttgac	gagcattttct	1380
aggggttcc	ccctcttcgc	aaaaaaggat	caaggatgt	tgaatgtcg	gaaggaaacg	1440
gttccttcgg	aaatgttcgt	aaagacaaaac	acgtctgtcg	cgaccctttg	caggcagccg	1500
aaccccccac	ctggcgacag	tgctcttcg	ggccaaaaagg	cacgtgtata	agatacaccc	1560
gcaaaaggccg	cacaaacccca	gtgcacccgt	ttgtgttgg	ttgtgttgg	aaggatccaa	1620
ttggctctct	caagctgtt	caacaaggg	ctgaaggatg	ccccaaagg	acccctatgt	1680
atgggatcc	atctggggcc	ttcggtggac	tgcttttacat	tggttttgc	gagggtttaaa	1740
aacgttcagg	ccccccggac	cacggggac	ttgtttttt	tttgcataaa	cgataataacc	1800
atggcgccata	ttacggccata	ctcccaacag	acgcgcggcc	tactttggctg	catcatact	1860

agccctcacag	ggccggacacg	gaaccaggcgt	gaggggggagg	tccaaagtgg	ctccaccgc	1920
acacaatttt	tcttggcgac	ctgcgtcaat	ggcgtgtgt	ggactgtcta	tcatggcgcc	1980
ggctcaaaga	cccttgcgg	cccaaaggc	ccaaatcaccc	aatgtacac	caatgtggac	2040
caggacatcg	tccgtgtcc	agggcccccc	gggggggggg	ccttgcac	atgcacccgtc	2100
ggcactgtcg	acctttaact	gttcacaggc	catggcgat	tcattccgg	gcccgggggg	2160
ggcgcacaca	ggggggggcc	actctccccc	aggcccggt	cttacttgaa	gggttcttcg	2220
gggggttcac	tgtctggccc	ctggggccac	gtctgtggca	tctttcggg	tgccgtgtgc	2280
acccgggggg	tttogaaggc	ggtggacttt	gtaccgtcg	agtctatgg	aaccactatg	2340
cggttccccc	tcttcaccca	caactcgcc	cttcggccgg	tacccgacac	atccaggatgt	2400
gccccatcat	acgccccatc	taatggccgg	aaagacacta	agggtccggc	tgctgtatgc	2460
gccaagggt	ataagggtgt	tgtctgtaa	ccgtcgtcg	ccgcacccct	aggtttccgg	2520
gcttatatgt	cttaacggaca	tgttgcac	cctaacaatc	gaacccgggt	aaggaccatc	2580
accacgggtt	ccccatcatc	gtactccac	tatggcaact	tcttcggcga	cgggtgttgc	2640
tctggggggc	cctatgacat	ataataatgt	gtatgtgc	actcaactga	ctcgaccatc	2700
atccatggca	tccggcacat	cttggaccaa	ggggacacgg	cttggacggc	actctgtgt	2760
ctccacccat	ctacgcctcc	gggatcggtc	accgtgcac	atccaaacat	cgaggagggt	2820
gcttgcac	gcaactgtgg	ataccccc	tatggcaac	ccatccccc	cgagaccatc	2880
aagggggggg	ggcacccat	tttgcgtcat	tccaaagaag	aatgtatgt	gctccggcg	2940
aagctgtcg	gcttgcgtact	caatgtgt	gtatattacc	ggggccctgt	tgtatccgtc	3000
ataaccaata	ggggagacat	tgttgcgt	gtcaacggc	cttataatgc	gggttttacc	3060
ggcattttcg	actcagtgtat	cgactgtca	acatgtgtca	cccaagacat	cgacttcac	3120
ctggccaccc	ccttcgtcat	tgagacacg	acccgtccac	aaagccgggt	gtcaacgtcg	3180
cagccggccg	gcaaggactgg	ttggggccgg	atggggat	acaggtttgt	gactccaggag	3240
gaacggccct	cgggcatgtt	cgattctcg	tttctgtgg	agtgtatgt	cgccgggtgt	3300
gcttgcgtac	agtcacccgc	ccggcagacc	tcaatgtgt	tcggggctt	cctaacaaca	3360
ccagggtttt	ccgttgcgtca	ggaccatctt	gaggtttgg	agagcgtt	tacaggcctc	3420
accacccatag	acgccccatc	tttgcgtcc	actaaacagg	caggagacaa	cttcccttac	3480
ctggtagcat	accaggatoc	gggttgcgg	agggttcgg	cttccaccc	atcggtggac	3540
caaatagtgt	agtgtctcat	acggcttaag	cctacgtgc	acggggccac	gccccctgtg	3600
tataggtctt	gagccatgtt	aaacagggtt	actaccac	cccccataac	caaataacatc	3660
atggatgtac	tgtccgtgt	cttggggat	gtcaacggca	cttgggtgt	ggtagggggt	3720
gtcttagcag	cttgcggccg	gtatgtgt	aaacacggca	gggtgtgt	tgtggccgg	3780
atccatgtt	cgggaaatcc	ggccatatt	ccoggacagg	aagtcttca	ccggggatgt	3840
gatgagatgg	aaagatgtgc	cttacaccc	catttacatcg	aacagggaat	gcagctcgcc	3900
gaacaatcca	aaacaaacggc	aaatccgggt	tttgcacaa	cccaacaa	agccggagggt	3960
gctgtccccc	tgttggatca	caatgtggcc	acccgtca	cttctgggg	gaagcatatg	4020
tggaaatttc	tcagcggtat	acaatattta	gtcggatgt	ccactctgtc	tggaaaccccc	4080
gctgatagat	cactgtgtgc	atttcaccc	ttatccaca	ccgcgtc	caccaacatc	4140
accctcttgt	ttaacatct	ggggggatgt	tttgcggcc	aacttgc	tcccgacgtct	4200
gctttctgtt	tctgtgtggc	ccggcatgtcg	ggggggatgt	tttgcggat	aggcccttggg	4260
aagggtgttg	tgtatatttt	ggcaggatgt	tttgcgggggg	tttgcgggggg	gtctgtggcc	4320
ttaaagggtca	tgagcggcga	gtatgcctcc	accggaggac	tgggttaa	actccctgt	4380
atccctcc	tttgcggccct	agtcgtcg	tttgcgtgt	ccatgtatgt	ggctcgccgc	4440
gtggcccccag	ggggggggcc	tttgcgtatgg	atggacccgg	tatgtatgt	cgcttcggcc	4500
ggtaaccaacg	tctccccccac	gtcaactatgt	tttgcgtatgt	ccatccccc	taacgtatgt	4560
cagatctct	tatgttttac	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	4620
gaggactgtt	ccacggccat	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	4680
acgggtgttg	tgtatgtttt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	4740
gtcccccttc	tcttcatgtca	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	4800
caaaccacat	ggccatgtgg	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	4860
atctgtggcc	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	4920
accacgggg	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	4980
gctgtgtgg	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5040
accactgtca	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5100
gatgggggtc	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5160
acatttcgtt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5220
ccggacatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5280
gtcaatgtca	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5340
ctgtgtggc	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5400
ctcatcgagg	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5460
tcagaaaaata	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5520
agggaaatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5580
cccatatgg	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5640
tacgttcctc	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5700
cctccacccg	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5760
gagctgcaca	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	tttgcgtatgt	5820

acggcccttc	ctgaccagcc	ctccgacgac	ggcgacgggg	gatccgaaegt	tgagtctgac	5880
tccctcatgc	cccccccttga	gggggaggccg	ggggatcccg	atctcagoga	cgggtcttgg	5940
tctacgtaa	gcgaggaggg	tagtgaggac	gtctgtctgt	gtctegatgtc	ctacacatgg	6000
acaggccccc	tgatcacgca	atcgccgtgg	gaggaaacca	agtcggccat	caatgcactg	6060
agcaacttt	tgtccctgtc	ccacaacttgc	gtctatgtca	caacatctcg	cagcgcaacg	6120
ctggccgaga	agaagggtcac	ctttgacaga	ctgcagggtc	tggaacgcca	ctacccggacg	6180
gtgctcaagg	agatgaaggc	gaaggcgctcc	acagttaaagg	ctaaaacttct	atccgtggag	6240
gaagcctgtg	agtcgtacgco	ccccacattcg	gocagatcta	aatttggcta	tggggcaaaag	6300
gacgtccgga	acctatccatcg	caaggccgtt	aaccatccatc	gtctccgttg	gaaggactgt	6360
ctggaaagaca	ctgagacacca	aatttgacacc	accatcatcg	aaaaaatgtg	gggtttctgc	6420
gtccaaccag	agaagggggg	ccgcaagcca	gtctgcctt	tcgttattccc	agattttgggg	6480
gttctgttgt	gcgagaaaaat	ggcccttttac	gtatgttgc	ccacccctccc	tcagggcggt	6540
atggcttctt	catacggtt	ccaaatcttc	cttgcacacg	gggttgcgtt	ctctgtgaat	6600
gcctggaaag	cgaagaatg	ccctatggc	ttcgatcatg	acacccgtgt	ttttgtgtgt	6660
acggcttctt	agatgtacat	ccgtgttgc	gagtcataat	accaatgttg	tgacttggcc	6720
ccccaaagcc	gacaggccatc	aaagggtcg	acagagccgc	tttacatcg	ggggccctcg	6780
actaattttc	aaaggccgaaat	ctgcggctat	ccggcggtcc	ggcggcgccg	tgtaactgacg	6840
accagctgtc	gttaatccatc	catgttgc	tttgcggac	ctggggctgc	tcgagctgog	6900
aaagtcctagg	actgcacatg	gtctgtatgc	ggagacgacc	tttgcgttat	ctgtgaaacg	6960
ggggggggaa	agaggacgg	ggcgacgact	cgggccttca	cggggttcat	gactatgatc	7020
tctgcccccc	ctggggaccc	ggccaaacca	gaatacgtact	ttggatgtat	aacatcatgc	7080
tcctccaaatg	ttgtacttgc	gcacgtatgc	tttgcggaaa	gggtgtacta	tctcccccgt	7140
gaccggccaa	cccccccttg	gggggtgtcc	ttggagacac	ctagacacac	tccaggtaat	7200
tcctggtag	gcaacatcat	catgtatgc	ccccacccgt	gggcaaggat	gatctgtatg	7260
actcattttt	tctccatctt	tttagtgc	tttgcggat	aaaaacccct	agattgtcag	7320
atctacgggg	cctgttactt	cattgcaccc	tttgcactca	ctcgcataat	tcaacgactc	7380
catggcccta	ggccatcttc	actccatagt	ttacttcgt	gttgcataat	ttagggtgtt	7440
tcatgtcc	ggaaacttgg	ggtacccggcc	tttgcgtatc	gggagacatc	ggccagaagt	7500
gtcccggtgt	ggctactgtc	ccagggggggg	agggtgtcca	tttgcggca	gtaccccttc	7560
aacttggggcag	taaggccaaat	gtccaaatcc	tttgcgtatc	ggctgtcg	ccagggtgtat	7620
ttatccatgt	ggttctgttc	ttttgttacacg	ggggggagac	tatatacaca	ctctgtctgt	7680
gccccacccc	gctgtgttac	ttttgttgc	tttttttttt	tttttttttt	tttttttttt	7740
ctactcccc	acccatgtac	ggggacttta	tttttttttt	tttttttttt	tttttttttt	7800
ttttttttttt	ttttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	7860
ttttttttttt	ttttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	7920
tagctgttgc	agggtctgtc	ggccgttgc	tttttttttt	tttttttttt	tttttttttt	7980
agatcaatgt	ctctctgtcc	ggccgttgc	tttttttttt	tttttttttt	tttttttttt	8040
gcatatgtac	ggccatccatc	ttatgttgc	tttttttttt	tttttttttt	tttttttttt	8100
agtttatcac	agttttttatgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8160
ctcatgtca	ttctccggcc	cggtttttttt	tttttttttt	tttttttttt	tttttttttt	8220
gtacttggcc	ggctttttttt	gtatgtatgc	tttttttttt	tttttttttt	tttttttttt	8280
gtgtgtgtgt	cgcttatatgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8340
tcccgacgtt	tttggccgtcc	ccaggatctgt	tttttttttt	tttttttttt	tttttttttt	8400
taoggatgtat	tttggccgtcc	tttgcgttgc	tttttttttt	tttttttttt	tttttttttt	8460
ggccatccagg	ggccatccagg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8520
gaagatgtgg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8580
ggcccccgttgc	ccggggggact	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8640
gcccgttgc	acggccatccaa	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8700
ggagacgtcc	ggccatgtatc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8760
cgggccatgtatc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8820
cagggttgcgg	cagggttgcgt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8880
atgtatccgg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	8940
actgtgtccgg	ccacccaaatcg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9000
gacggcgcttc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9060
atgatattttc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9120
caggatgtatc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9180
acttcgtatca	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9240
aacgggtttgg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9300
egggggccat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9360
gatttacccat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9420
ccaaaccccttg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9480
tctccggggcc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9540
ccgggtttagg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9600
gaacgttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9660
tcgggtttcc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9720
tcggatgtttc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	9780

cgaaaggcgtg	gcattggaccc	tgagtgttatt	ttctctgtgtc	cggccgcata	cataaccgcac	9840
gttgtttaacc	ctcacaacaat	tccagtaacc	gggcattttc	atcatcagtc	accgcgtatcg	9900
tgagcatctt	cttcgttgc	atccgtatca	ttacccccc	caacagaaat	tcccccttac	9960
acggaggcat	caatgtggcca	aacaggaaaa	aacccgcctt	aacatggccc	gctttatcg	10020
aaggccagaca	ttaaacgcctt	tggagaaaat	caadcgatgt	gacgcggatg	aacaggcaga	10080
catctgtggaa	tcgttgcac	accacgcgtg	tgagtttac	cgccgcgtcc	tcgcgcgtt	10140
cgtgtatgac	gttgcggaaaat	tgcacat	gcgcgttcc	gagacgttcc	cagttgtct	10200
gttaagcgat	gcccggggca	gacaaggcccg	tcaggcccg	tcagccgggt	ttgcgggggt	10260
tcggggggca	gcoatggacc	agtccatgt	cgatagcgg	gtgtataact	gtttaactat	10320
gcccgcattat	cgaggatgtt	actggagatgt	cacatatgc	ggtgtgaat	acacgcacaga	10380
tggtgtaaagg	gaaaataccg	catcgccgg	tcttcgttcc	cctgcgttac	tgactcgctg	10440
cgctgggtcg	tcgggtcg	gagacgttgc	tcagttact	caacaggcgtt	aatacgttta	10500
tccacagatc	cagggtatata	cgcaggaaag	aacatgttg	caaaaggccca	gcaaaaaggcc	10560
aggaacgcgtt	aaaaaggcccg	gttgcgtggg	ttttccat	ggttccggcc	ccctgtacgg	10620
catcacaaaaat	atcgacgttcc	aatgcagagg	ttggggaaaacc	cgacggactt	ataaaagatcc	10680
caggcgttcc	ccccctggaa	ctccctctgt	cgcttccgt	ttccgcacdt	gcccgttacc	10740
ggatatactgt	cccgcttcc	cccttcgttgc	agegtggcc	tttctcatag	ctcacgtgt	10800
aggatcttc	gttgcgttgc	gttgcgttgc	ttcaacgttg	gctgtgtc	cgaaacccccc	10860
gttcagcccg	acgcgttgcg	tttacccgtt	aatactatgc	tttgcgttcc	ccccgtttaaga	10920
cacgactttt	cgccacttgc	agcagacactt	ggtaacaggcc	tttagcagac	gaggatgtat	10980
ggcggtgtca	cgaggttctt	gaagtgttgg	cotaactacg	gttacactag	aaggacacgtt	11040
tttgcgtatct	ggcgttgcgt	gttgcgttgc	acccctcgaa	aaaggttgg	tagcttttgc	11100
tccggcaaac	aaaccacccgc	tggtagacgtt	ggtttttttt	tttgcgttgc	gcaaggatccg	11160
cgcagaaaaaa	aaaggatctca	agaatgttcc	tttgcgttcc	ctacggggcc	tgacgttccgt	11220
tggaaacgaaa	actcgttcc	agggttttgc	gttgcgttgc	tatcataaaag	gattttcacc	11280
tagatctttt	tcttagataat	acgactcaact	ata			11313

<210> 3
 <211> 11313
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid

<400> 3						
gcccggccccc	gattgggggc	gacactccac	catagatcac	tcccccgtgt	ggaactactgt	60
tcttcacggc	gaaaggccgt	atccgttgcgg	tttagtatgt	tggtgtc	cttcuaggac	120
ccccccctcc	ggggagaccc	tagtgcgtt	cggaacccgt	gagttacac	gaatttgcac	180
gacgacccgg	ttcttttgc	gttcaaccccg	ttcaatgcgt	ggagattttg	gcgtgcggcc	240
gcgagatgc	tagccgtat	gttgcgttgc	gaccaaggcc	ttgtgttact	gcctgtatagg	300
gtgtgtgg	gttgcggccgg	aggctctgt	ggccgttgcac	catgagcacc	aatccaaac	360
ctcaaagaaa	aaaaacaaagg	ccggccatgt	tttgcgttgc	tttgcgttgc	ccagggttcc	420
ccggccgtt	gttgcgttgc	tttgcgttgc	atccgttgc	tttgcgttgc	atccgttgc	480
ctgtatccgc	cggttccgg	tttgcgttgc	aggggccccc	ggttttttt	gtcaaggaccc	540
acctgtccgg	tcgccttgc	gttgcgttgc	aggccggcc	tttgcgttgc	tttgcgttgc	600
cgaggccgg	tccttgcgt	gttgcgttgc	acgttgcgt	tgaaggccgg	aggggactggc	660
tggtatgttgc	cgaaatgcgt	ggggccgttgc	tttgcgttgc	tttgcgttgc	cttgcgttgc	720
aatgtatcc	catgtgttgc	gttgcgttgc	gggttgcata	gttgcgttgc	gttgcgttgc	780
catttgcgtt	ccaaacggaaa	atccgtatcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	840
tttgcgttgc	ggatgtatcc	gaccaatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	900
ccaggcttac	ggggccgtat	ccggccggcc	aggatctgt	ctgttccat	ggcgtatgt	960
gttttgcgtt	tatcatgttgc	gaaatgttgc	gttgcgttgc	tttgcgttgc	tttgcgttgc	1020
tgggtgttgc	ggacgttcc	caggatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1080
tttgcgttgc	atgggtgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1140
agccgtatcc	tttccatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1200
tttgcgttgc	atccgtatcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1260
cgaaaggccgt	tggaaataagg	ccgggtgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1320
ccgttccat	ggcaatgttgc	ggccggccaa	tttgcgttgc	tttgcgttgc	tttgcgttgc	1380
agggttgc	cccttccat	caaaaggatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1440
gttgcgttgc	aaaggatgttgc	aaaggatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1500
aaccccccac	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1560
gcaaaaggccgt	cacaacccca	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1620
tttgcgttgc	caacaaaggccgt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1680

tacgtccctc	cagtggtaaca	cggtgttcca	ttggccgcgt	ccaaggcccc	tcgatatacc	5700
cctccacggg	ggaaagggac	gggtgtctcg	tcagaactta	cgcgtgttcc	tgcccttggcc	5760
gagctcgcca	caaagacctt	cggcagctcc	gaatcgctgg	cgcgcacag	ccgcacggca	5820
acggcccttc	ctgaccggcc	cttccgacac	ggcgcacggg	gatccgcgat	tgatgtcgac	5880
tcctccatgc	cccccccttg	ggggggggcc	ggggatccct	atctcagoga	cgggtcttgg	5940
tctacatggc	gggggggggg	ttttttttgg	ttttttttgg	ttttttttgg	ttttttttgg	6000
acaggcgccc	tgatcagccc	atgcgttgcg	gaggaaacca	agctgcctcc	caatcgactg	6060
agcaactctt	tgctccgtca	ccacaacttg	gtctatgtca	caacatctcg	cagcgcacagc	6120
ctggggccgg	agaaggatcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6180
gtgtcaagg	agatgggggg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6240
gaagccgtta	agtttttttt	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6300
gacgtccggg	acccatccatcg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6360
ctggagacca	ctggagaccc	aatttgacacc	accatcatgg	aaaaaaatgg	ggtttttctgc	6420
gtccaaaccg	agaagggggg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6480
gttgcgttgt	ggggaaaat	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6540
atgggccttt	catacgattt	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6600
gcctggaaag	cgaaagaaatg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6660
acggctactg	agaatgtacat	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6720
cccgaaacca	gacggccat	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6780
actaatitca	aaggggggaa	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6840
accagctgg	gttaataccct	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6900
aaacctccgg	actggcagat	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	6960
ggggggccgg	aaaggaggcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7020
tctgcccccc	ctggggaccc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7080
tecccaatg	tgtccatgc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7140
gaccggccca	cccccccttg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7200
tcctggccatg	ggcaatccat	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7260
actcattttt	tctccatctt	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7320
atctacgggg	ctgtttaact	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7380
catggcccta	ggccatctttt	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7440
tcatgcctca	ggaaactttt	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7500
gtccggccat	ggctatgtcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7560
aactgggggg	taaggggccaa	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7620
ttatocagct	gggttcgttc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7680
gccccggccc	gggtgttcat	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7740
ctactcccca	acccatggatc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7800
tttccctttt	tttttttttt	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7860
tttttctct	tttttttttt	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7920
tagctgtaa	agggtcgttgc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	7980
agatacgta	cttcgtcagg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8040
gcatatttgc	gaccggatcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8100
agttttatca	agtttaaatgg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8160
ctcatgttca	ttccctggcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8220
gtactggccg	ggccctttcg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8280
gtgtctgtat	cgctatgtcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8340
tccgacccat	ttggccggcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8400
tacggatata	ttggccggcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8460
ggccatccgg	ggccggccgg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8520
gaagatccgg	ctggccatcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8580
ggcccccgtt	ccgggggggg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8640
goggtgttca	acggccatca	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8700
ggagacgtcc	gaccggatcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8760
cgggggatcg	ctatcgctcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8820
cagggttccgg	caggcgctcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8880
atgatccgg	tgtcgcttgc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	8940
actgttccgg	ccacccaaatgg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9000
gaoggccgtt	gtcaatgttcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9060
atgatccatc	tcgtcgatccgg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9120
caggtagatg	acgaccatcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9180
acttcgtatca	ctggacccgt	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9240
aacgggttgg	catgttgttcc	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9300
cgcgggtat	gggggggggg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9360
gattcacccat	tccaaatgg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9420
ccaacccctt	ggcaaaatgg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9480
tctoggccat	ctgggggggg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9540
ccgggtttaggg	ctgggggggg	tttttgcacca	tttttgcacca	tttttgcacca	tttttgcacca	9600

gaacgttgaag	cgtactgtgc	tgcacaaacgt	ctgcgcacgt	agaacaaca	tgaaatggct	9660
tccggtttccg	tgtttcgtaa	agtctggaaa	cgccggaaatc	accattatgt	9720	
tccggatctgc	cattcgagg	tgcgtgttgc	atccatgttc	tctgttatgg	9780	
cgaaggcgtc	gtattgaccc	ttagtgattt	ttctctgttc	aacaccatca	9840	
gttgttattcc	ctcacaacgt	tccatgttacc	ggcgcgttcc	cgtccgcata	9900	
tgacatccct	tctcggttcc	atccgttatac	atcatcgatc	accctgtatcg	9960	
acggaggccat	caatgtgacca	aacaggaaaa	atccccccat	gaacaaatgt	10020	
acccggacaga	ttaaaggctt	tggagaaaaact	aaccggccctt	aatcatggccc	10080	
cattgtgaa	tgcgttcaacg	accacgctga	caacaggttc	gacccgttgc	10140	
cgggtatgtac	ggtgaaaaacc	tctggacat	ttaggttttc	cgccagtttc	10200	
tgaaaggcgat	ggggggggca	gaaaccccg	ttgggggttc	tgccgtgttgc	10260	
tcgggggcgc	gcccattgtt	agtcacgtag	ttgggggttc	ttgggggttc	10320	
ggggccatcg	agcagatgtt	actggagatc	gttgcataatc	gttgcataatc	10380	
tgccgttgg	ggaaaaatccy	cattcgccgc	caccatatgc	gggtgtttat	10440	
tcaccaaat	tgcgttgcgg	ggggggggta	tcttcgttgc	cctccgttgc	10500	
aggaaacgta	caggggatata	cgccaggaaa	tcatgcgtact	tgactcgctg	10560	
catccaaaaa	aaaaggccgc	gttgcgtggc	caaaaggccgt	atatacggtt	10620	
cagggtttc	atccatgcgtc	ttttccatca	ggccggccccc	ggccggccccc	10680	
ggataccctgt	ccccctggaa	ttttccatcg	ggctccgttc	ccatggactac	10740	
aggatcttc	cgcgtttttt	cccttcgtgc	tttccgcattt	ggccgttacc	10800	
gttcaccccg	gttcgggtta	gttcgttgc	ttttccatcg	cttcacgtgt	10860	
cacgattat	acccgttgc	tttccatcg	tttccatcg	tttccatcg	10920	
gggggtgtca	cgccatgtt	aggacggact	aactatcg	ttccatcg	10980	
tttggatctt	cgatgttttt	aaatgtgttg	ggtaacaggaa	ttccatcg	11040	
tccggcaaac	ggcttttttt	ttttccatcg	tttccatcg	tttccatcg	11100	
cgcggaaaaa	aaaccacccgc	ttttccatcg	tttccatcg	tttccatcg	11160	
tggaaacggaa	aaaggatctt	aaagatgtt	tttccatcg	tttccatcg	11220	
tatcatcttt	actccatgtt	agggtttttt	tttccatcg	tttccatcg	11280	
	ttagatgtat	tttccatcg	tttccatcg	tttccatcg	11313	

```
<210> 4
<211> 11313
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Plasmid

<400> 4

ggccggcccc	gattgggggc	gacactccac	catagatcac	tcccctgtga	ggaactacty	60
tcttcacgca	gaaggactgtt	accgttgcgg	tttagatcgag	ttcgctgtcg	cctccaggac	120
ccccccctcc	ggggagacca	tagttgttcg	ggaaacccgg	gagatacccg	gaattgcag	180
gacgaccggg	tcttttcttg	gatcaaccgg	ctcaatgcct	ggagatttgg	ggctgtcccc	240
ggcgactgtc	tagccggata	tggtgggttc	ggcaaaaggcc	ttgttgtact	gtgtatgat	300
tggtgtcgca	tgccccccgg	agtgttcgtc	gaccgttcgg	catgagacgg	aatcctaaac	360
ctcaaagaaa	aacccaaaggg	cgccgcgtat	ttgaacaaggc	ttgtatgcac	gcagggttctc	420
ccggccgttt	gttggagagg	ttatccgtct	atgactggcc	acaacagaca	atcgctgtct	480
ctgtatggcc	ctgtttccgg	ctgttcacggc	aggggggccgc	gttctttttt	gtcaagaccg	540
acctgttcgg	tgccctgtat	gaactgcggg	ggggcggcagc	ggggctatcg	ttgtgtccca	600
cgacggggct	tccttgcgca	gttgcgttcg	acggttgttc	taaggccggg	aggggactggc	660
tgttatgtgg	qaaggatgcg	ggggcggatc	tcctgttcac	tcacccgtct	cctgcggcaga	720
aagtatccat	catggctgtat	ggctgcataac	ggctgtatcg	gttactatggc	gttactatggc	780
cattcgacca	ccaaaggaaaa	catcgatcg	tactcgatgg	aaaggccgtc	aaactgttgc	840
tttgcgtatc	ggatgatctg	gacggaaagac	atcaggggct	cgccggccacg	aaactgttgc	900
ccaggcttca	ggccggcgtat	ccggccgggg	aggatcttcgt	ctgttgcggat	ggcgatgtct	960
gttgcggcaa	tatcatgtgt	ggaaaatggcc	ttttttttcg	attatcgatc	ttgtggccgc	1020
ttgggtgtgg	ggaccgtatc	caggacatag	ttttttttcg	ctcgatgtatc	gttgaagagc	1080
tttgcggcga	atgggtgtac	cggttcttcgt	ttttttttcg	ttttttttcg	ttttttttcg	1140
agccgtatcc	ctttatcgct	ttttttttcg	ttttttttcg	ttttttttcg	ttttttttcg	1200
gtttcccttc	atggccgtatc	ttttttttcg	ttttttttcg	ttttttttcg	ttttttttcg	1260
cgaaggccgt	tggataaagg	ccgggtgtgg	ttttttttcg	ttttttttcg	ttttttttcg	1320
cgcgtttttt	ggcaatgtgg	ggccggggaa	cttggccctg	ttttttttcg	ttttttttcg	1380
agggggtttttt	cccccttcgtc	caaaaggatg	aaacgttctgt	ttttttttcg	ttttttttcg	1440
atcccttcgt	aaggcttttc	aaaaacaaaca	acgttctgtatq	cccccttcgt	cccccttcgt	1500

aaccccccac	ctggcgacag	gtgccttcgtc	ggccaaaagg	cacgtgtata	agatcacact	1560
gcaaggcgg	cacaaccccc	gtgccccgtt	gtgagtggaa	tagttgtga	aagagtcaaa	1620
tggcttct	caagcgatt	caaaagggg	ctgaaaggat	cccagaagg	acccttctgt	1680
atgggatct	atctggggcc	tcggtcaca	tgcgttatcat	gtgtttatgc	gagggtaaaa	1740
aacgtctagg	ccccccggaa	cacggggacg	tggttttctt	tgaaaaaaca	cgataatacc	1800
atggccctta	ttacggccct	ctcccaacag	acgcggaggc	tacttgggtc	catcatcact	1860
agcgtcacag	ggccggacag	gaaaccaggc	gagggggggg	tccaaqtgtt	ctccaccqca	1920
acacaatctt	tccggcgac	ctgegtcaat	ggcgtgttt	ggactgtcta	tcatgtgtcc	1980
ggcttcaaaa	cccttgcggg	ccaaaaggcc	ccaaatccacc	aaatgtacac	caatgtggac	2040
caggacctcg	tcggctggca	aggccccccc	gggggggggt	ccttgacacc	atgcaccatcg	2100
ggcagcgggg	acctttactt	ggtcaagggg	catggccatg	tccatcccggt	gcccgggggg	2160
ggcggaccca	ggggggagct	acttcctccc	aggccgggtt	cctacttgaa	gggttcttcg	2220
ggcgttccac	tgcttcggcc	ctcggggac	gctgtgggca	tctttcggtc	tgccgtgtgc	2280
acccgggggg	tttgcggggc	gggtggactt	tgaccctcg	actctatggg	aaccactatg	2340
cggccccccc	tcttcgggg	caactctgtc	cctccggggc	tacccggacag	atcccgagggt	2400
gccccatctac	acgccttac	tggtagccgc	aagagacta	agggtgcggc	tgcgatgtca	2460
gccccaaagggt	ataaagggtc	tgcttcgaad	cogtcgtcg	ccgcacccct	agggttccgg	2520
ggtatataat	ctaaaggccat	tggtagatcg	cctaacaata	gaacogggggt	aaggaccatc	2580
accacgggtt	ccccccatcc	gtactccacc	atgtggcaat	tttttgcgca	cggttgcgtc	2640
tctggggggc	ctatgtacat	cataaatatg	gtatggatgc	actcaactgt	ctcgaccatc	2700
atcttgggca	tcggcacagt	cttggaccaa	gccccggagg	cttggacgcgc	actctgtgt	2760
ctggcccaacg	ctacgcgtcc	gggatggctc	acggtcgcac	atccaaacat	cgaggagggt	2820
gctctgtcgc	gcaactgggg	aatccccctt	tatggcaaaa	ccatccccat	cgagaccatc	2880
aaggggggga	ggcacccat	ttttcgccat	tccaaagaaa	aatgtatgt	gctccggccgg	2940
aagctgtccg	gctctggact	caatgtgtca	gcatattacc	ggggcccttg	tgtatccgtc	3000
ataccaaact	ggcgagacgt	catgtgtca	gcaacggagg	ctctaataatg	gggtttaacc	3060
ggcgatttcg	actctgtgt	cgactgtcaat	atcatgtgc	ccccagacatg	cgacttccgc	3120
ctggaccctt	ccttcacat	tgagacggcc	acccgtgcac	aaagccgggt	gtcactcgct	3180
cagcggcgag	gcaggactgg	tagggggagg	atggccattt	acgggtttgt	gactccaggag	3240
gaacggccct	ggggcatgtt	cgatccctcg	gtttctgtcg	agtgctatgt	cgccggctgt	3300
gctttgtatag	agcttcaccc	cgccggacatc	tragttaggt	tgccgggttt	cctaaacaca	3360
ccagggttgc	ctggctgcac	ggaccatctg	gatgttggg	agaggcttgtt	tacaggccctc	3420
accacatcat	agcggccatt	cttgcaccc	actaaaggcc	caggagacaaa	cttccctatc	3480
ctggtagat	accaggctac	gggtgtcgcc	agggtctagg	cttccaccc	atcggtggac	3540
caaataatgg	agtgttctat	acggcttaag	cttacgttcg	acggggccac	gccccctgt	3600
tataggcttc	ggaggcttca	aaacggaggat	actaccacac	acccctataaa	caataacatc	3660
atggcatca	tgtcaaggta	cctggggatc	tcacggacga	ccttgggtgt	ggtagggcgg	3720
gtccatcgat	cttcggccgc	gtattgtcg	acaacaggac	gggtgtcat	tgtggccagg	3780
atcatcttt	cgccggaaat	ccggccatcat	cccgacagggg	aagtctttta	ccggggacttc	3840
gatgatgttgc	aaagggtggc	ctcacactc	cttacatcgt	aacggggaaat	cgacgtgtgc	3900
gaacatttca	baacagggaa	aatccgggtt	ctgcaaaacag	ccaccaaggaa	agccggagggt	3960
gctgtccccc	cggttggaaatc	caagtgggg	accctcgaa	ccttctggcc	gaagcatatg	4020
tggatatttc	tcacggggat	accaatata	cgaggctgt	ccactctgc	tggcaacccc	4080
gctgtatgc	caactgtatgc	attcacagcc	ttatcatccaa	gcccgtctca	cacccaaacat	4140
accctcttgt	tttaatacttc	ggggggatgtt	tttgcggccccc	aactgtcttc	tcccacggcc	4200
gttttgcgtt	tgttagggcc	cgccatgtatc	ggaggccgttt	tttgcggatgt	aggcccttgg	4260
aagggtttgt	tggatatttt	ggcagggttat	ggagcaggggg	ttgcaggcgc	gtctgtggcc	4320
ttaaaggta	ttggggggatc	gtatccccc	acccggggcc	ttgtttaacct	actccctgt	4380
atccctcttc	cttggggccc	actgttcggg	tttgcgttgc	caggcatatc	ggcttggccat	4440
gtggggccat	gggggggggg	tgttcgttgc	atgaacccgc	tttgcgtatgt	egcttcgggg	4500
ggttaaaccac	tcttcccccc	gcatatgttc	cttgcggatgc	agctgtcgac	atcggtcaat	4560
catagcttc	cttagtcttc	catcaactca	ctgtgtgaaga	ggcttcacca	ttggatcaac	4620
gaggactgt	ccacggccat	ctccgggtcg	ttggctaaagg	atgtttggaa	ttggatatgc	4680
acgggtttgg	ctgtatttca	gacccgttc	cagtcccaat	ttctccggcc	attccgggggg	4740
gtcccccttc	tctcatgtca	acgtgggtac	aaggaggatc	ggccgggggg	ccggcatatcg	4800
caaaccac	ggccatgtgg	agccacatgc	acccggatcg	ttccatgggg	ttccatggagg	4860
atctgtggcc	cttggacccat	tagtaacatc	ttggatcgaa	cattccatcc	taacgcgtac	4920
accacggccc	cttcggccgc	cttccggccg	ccaaatattt	tttggggctt	tttgggggtgt	4980
gctgtgttgc	agatgttgc	gggttaacgg	tttggggatgt	tccactactgt	gacggccatgc	5040
accactgaca	acgttaaaatg	cccggttcag	tttggggatgt	ccggatattt	cacagaatgt	5100
gatgtgggtc	gggttgcacag	gtatccgtca	gggtgtcaac	cccttcatacg	ggaggaggatc	5160
acattcttc	tcggcttcgc	tcataacttc	ttttgggttc	acgtccatcat	cgaggccggaa	5220
ccggacatgt	cagtgtccat	ttccatgttc	acccggccat	ccacccat	ggccgggggg	5280
gctaagccat	gggttgcggcc	gggtatccctt	ccctcccttg	coagcttcatc	agctatccatc	5340
ctgtctgtcc	tttccttgc	gaaacatcatc	actaccatgt	atgatcccc	ggacgtctgac	5400
ctcatcgagg	ccaaacccatct	gtggcgccag	gagatggggg	ggaacatcac	cccgctgggg	5460

caaccaccttgc	cgacaata	tccatcgctg	ccggccatctc	cagcaggccg	aacgcgggca	9480
tctccggccatgc	cgttgggttc	tggccacggg	tgcgttactatc	gttagcagaat	ctgttgaggaa	9540
ccccggcttagg	ctgggggggt	tgcctttatc	gtccggacccat	taaatcacccg	atacgcggac	9600
gaacctgttaag	caactgtctgc	tgcaccaaaatc	agcggggaaatc	acaacaaacaa	taatgtgttc	9660
tcgggtttccg	tgtttcgtaa	agtctggaaa	cgccggaaatc	cgccggccatc	accattatgtt	9720
tccggatctcg	catcgaggaa	tgcgtctggaa	tacccctgtgg	aacaccatcata	tctgttattaa	9780
cgaaaggctgtc	gtatggacc	tgtatggatcc	ttcttcgttgc	ccgcgcgatc	cataccggcca	9840
gttgtttacc	ctccaaacatc	tccgttaccc	gggcgtatc	atccatcgatc	accggatatcg	9900
ttagccatctc	ctctcggttc	atccgttatca	ttaccccccatt	aaacagaaaaat	ttcccccattac	9960
acggggccat	caatgtgacca	aaacaggaaaa	aaccggccat	aaacatggccc	gtttttatccat	10020
aaggccacat	ttaaatcgctt	tggggaaatct	caacgcgtgc	gacccggatgc	acaggccatgc	10080
cacatgtgtaa	tcggcttcacgg	accacccgtga	tgatgtttac	cgccagtcgc	tccgcgttgc	10140
cgggtatgtac	gtgtggaaacc	tgttgcacatc	gtatggccatc	gagacggcttc	cagctgttgc	10200
tgaaaggccat	ggggggagca	gaaacacccg	tcaggggcg	tcaggccgtg	ttgcgcgttg	10260
tcggggggccat	ggccatgtacc	agtccgtatg	cgataggccg	gttatactatc	gttataactatc	10320
ggggccatcg	agcagatgtt	atctggatgt	cataatccat	gtgtgtaaat	accggccacaga	10380
tcgtcaatgg	aaaaataccg	cataggcgcc	tcttcgttgc	cetcgcgtac	tgactcgctg	10440
cgctccgtgc	ttccgttgcg	ggcgaggctt	tcatgcgtact	caaaaggccgt	ataacgggtta	10500
tccacaaatgt	caggggataa	cgccaggaaaa	aaatctgttgc	gttgcgcggcc	gcaaaaggccc	10560
aggaaacccgt	aaaaggccgc	gttgcgtggg	ttttttccata	gttgcgcgc	ccctgcgacgg	10620
catccaaaaat	atcgacgtcc	aagtccaggg	ttttttccata	gcacaggact	ataaaatgtac	10680
caggccgttc	ccccctggaaat	cttcctcgatg	ttttttccata	tttcgcgcct	ggccgttacc	10740
ggatataatcg	ccggccctttt	cccttcgggg	acgtccgtgg	ttttttccata	cttcacgtgt	10800
aggatataatcg	gttcgtgtatc	gttcgttgcg	ttttttccata	gtgtgtggc	cgaaaaaccgg	10860
gttcacggcc	acggctgtgc	tttattccgt	aactatctgc	tttagtccaa	ccccggataga	10920
cacggatattat	ccggccatgg	acggccatgtc	gttaacaggaa	tttagcagage	ggaggatgtta	10980
gggggtgtctca	caggatgttt	ttttttccata	ttttttccata	ttttttccata	ttttttccata	11040
tttggatatcg	ttttttccata	ttttttccata	ttttttccata	ttttttccata	ttttttccata	11100
tccggccaaat	aaacccacccg	ttttttccata	ttttttccata	ttttttccata	ttttttccata	11160
cggccaaaaat	aaaggatctca	ttttttccata	ttttttccata	ttttttccata	ttttttccata	11220
ttggaaacccaa	attccacgtta	ttttttccata	ttttttccata	ttttttccata	ttttttccata	11280
taatcatctt	ttttttccata	ttttttccata	ttttttccata	ttttttccata	ttttttccata	11313

```
<210> 5  
<211> 11313  
<212> DNA  
<213> Artificial Sequence
```

<220>
<233> Plasmid

<400> 5

ggcagecccc	gattgggggc	gacactccac	catagatcac	tccctgtga	ggaactactg	60
ttttcacgca	gaaggatctt	aggccatggc	tttagtgcac	tgcgtcgacg	cctccaggad	120
ccccccctcc	ggggagacca	tagttgtctg	ggcaaccggg	gagataccgg	gaatttgcgg	180
gacgaccggg	tcctttctt	gatcaacccgg	ctcaatgcct	ggagattttg	gcgtgcccccc	240
gcgagactgc	tagccggata	tggttgggtc	ggcaaaaggcc	tttgtgttat	gctgtatggat	300
tggtttcgca	tgccccccgg	aggatccgtc	gaccgtcgac	catgagcagc	aatccctaaac	360
ctcaaagaaa	aaccaaaggg	cgccgcattg	ttgaacaaggc	tggatgtcgc	gcagggttctc	420
cgccgcgttgc	ggttggaggg	ttatccggct	atgactggcc	acaacagaca	ateggctgtct	480
ctgtatggcc	ctgtttccgg	ctgttcggcg	aggggggccc	gttctttttt	gtcaagacgg	540
accttgcggc	tgccctgtta	gaactgcggg	acggaggcgc	gcggctatcg	tggctggccca	600
cgacggcgct	tccttgcgca	gtgtgtctgg	aggatgttcg	taacggggcc	aggggactggc	660
tgatctatggg	cgaagtgcg	gggcggatgc	tcctgtatcc	tcaccttgt	cctgccccaga	720
aagtatccat	catggctgtat	gaaatcgccg	ggctgcatacc	gttgcattccy	gtcaacttgtc	780
catttcgcaca	ccaaaggcataa	catcgatcg	acgcggacgc	tactcgatgg	ggacccggcgc	840
ttgtcgatca	ggatgtatctg	gacggaaagac	atcaggggct	cgccgcaccc	gaacttgcgc	900
ccaggcttcaa	ggccgcgcatg	cccgacggcc	aggatctcg	ctggatccat	ggcgatgtcg	960
gtttttcgaa	tatcatcggt	aaaaatggcc	gtttttttcg	atccatcgac	tgttgcgggc	1020
ttgggtgtgc	ggaccgtatc	caggacatag	ctgttgcgtac	ctgtgtatcc	gttgcggatgc	1080
ttggccggca	atgggtgtac	cgtttccatcg	tggtttttcg	atccatcgcc	ttttttttccgg	1140
ttggccggca	tttttatcgca	ctttttatcg	atggttttcg	atggggccgg	ccggatccgc	1200
gtttttctct	atggggcccc	atccggcccc	ttttttttcc	atgtttttat	ccggatccgc	1260
ccaaaggccct	ttttatcgaa	ccggatccac	ttttttttcc	ttttttttat	ccggatccgc	1320

ccgtcttttgc	caatgttgaa	ggccccggaaa	cctggccctg	tcttttgcac	gaggattact	1380
agggttctttt	cccctctcg	caaggaaatg	caaggctctgt	tgaatgttgt	gaaggaaacg	1440
gttcccttgg	aaaccttttt	aaagacaaca	acgtctgtc	cgaccctttt	caggcagccg	1500
aaccctccca	ctggcgcac	gtgcttcgtc	ggccaaaagc	cacgtgtata	agatacacct	1560
gcaaaggccg	cacaacccca	gtgcccacgtt	gtgagttgga	tagttgtga	aagagtcaaa	1620
tggcttctct	caagcgattt	caacaagggg	ctgaaggatg	cccagaagg	acccccattgt	1680
atggatgttt	atctggggcc	tcgtgtcaca	tgttttacat	tggttttgtc	gaggtaaaaa	1740
aacgtctagg	cccccccgaa	cacggggaa	tggttttct	ttgaaaaaac	cgataatacc	1800
atggggccta	ttatcggtct	cttcccaac	acgcgaggac	tacttgggt	catcatcaet	1860
agcttcacag	ggccggacag	gaaccaggat	gaggggggac	tccaaagtgt	cttccacccgca	1920
acacaatctt	tctggggcag	ctcgctcaat	gggtgtgtt	ggactgttca	tcatgtgtcc	1980
ggcttaaaga	cccttgcgg	ccccaaagg	caatcacc	aatgttacac	caatgtggac	2040
caggactcg	tcggctgca	agccccc	ggggccgtt	ccttgcac	atgcacccgtc	2100
ggcagcttgg	actttttact	gttgcagg	catgcgtat	tatcccggt	ggccggccgg	2160
ggccgacacg	ggggggaccc	actctccccc	aggccctgt	ectacttgg	gggttcttgc	2220
ggcgttccac	tgcttcgtcc	ctggggggac	gttgcggca	tcttcgggg	tgccgttgc	2280
acccgagggg	ttgcgaagg	gggtggactt	gtaccctgt	agtcattatgt	aaccatctg	2340
cggtccccc	tcttcacgg	caactctgtc	cctccggcc	tacccggac	attccagggt	2400
gccccatctac	acggccatca	tgttagggc	aaagacact	atgtgcggc	tgctgtatgo	2460
gccccagggt	ataagggtct	tgttcgttac	ccgtccgtt	ccggccaccc	aggtttgggg	2520
gcttatatgt	ctaaggcaca	tgttgcgtac	cataacatca	gaacccgggt	aaggaccatc	2580
accacgggtt	ccccccatcc	gtactccacc	tatggcaat	ttcttcgcga	cggtgttgc	2640
tctgggggg	cctatgacat	cataatatgt	gtatgtgtt	actcaactgt	ctcgaccat	2700
atccctggcc	tcggcaactg	cctggggcc	ggggagacgg	actggagccg	actctgtgt	2760
ctgcacccac	ctagctcc	ggggatgtt	acccgttccac	atccaaacat	cgaggagggt	2820
gctctgtcca	gcactggaa	aatccccc	tatggcaaa	ccatccccc	cgagaccatc	2880
aagggggggg	ggccatccat	tttctgcgt	tuccaaga	aatgttgcata	gttcggccgg	2940
aaagtgttgc	gcctcggtat	caatgtgtt	gtatattac	ggggcccttgc	tgtatccgt	3000
ataccacta	ggcgaggatc	cattgtgtt	gcaacggac	cttcaatgc	gggttttacc	3060
ggcgatttgc	actctgttgc	cgactgtat	acatgtgtc	cccaagacat	cgacttcacg	3120
ctggaccctt	ccttcacccat	tgagacgac	accgtgcac	aaagccgggt	gtcacgttc	3180
cageggggag	ggcaggactt	ttggggcagg	atggccat	acagggttgc	gactccaggaa	3240
gaacggccctt	cgggcatgtt	cgatctccgt	gtttctgtgg	agttgtatgc	cgccgggttgc	3300
gcttggta	agtcacgc	cgccggaccc	tcagttgtt	tgccgggttca	ccttaaacaca	3360
ccagggttgc	ccgtctgtcc	ggacccatctg	gaggtttgc	agagcgttcc	tacaggcctc	3420
accacatag	acggccatcc	cttgcgtcc	actaaaggcc	caggagacaa	cttccctatc	3480
ctggatcat	accaggatct	gggtgtgc	gggtgtccagg	cttccatctc	atccatggcc	3540
caaaatgttca	aaacaaaggc	aatcggttgc	ctgcaac	ccaccaagca	gccccctgtc	3600
gctgttcccg	cggtggaaatc	caagtgggg	accctggaa	ccttgcgttgc	aaatatacatc	3660
tataggttgc	ggccgttca	aaacggatgt	actaccac	acccttataac	gaaatatacatc	3720
atggcgttgc	tgtcgttgc	cctgggggtt	gtcagcgac	ccttgggtgt	ggtagggccg	3780
gtcttagcag	cttcggccat	gtattgtcc	acaacaggac	gggtgttgc	tgtggggcagg	3840
atcatcttgc	ccggaaagcc	ggccatattt	cccccagg	aacttttttta	ccggggatcc	3900
gatgagatgg	aaagatgtgc	ctcacccatc	cattacatc	aacggggaaat	gcagcttcac	3960
gaacatttca	aaacaaaggc	aatcggttgc	ctgcaac	ccaccaagca	agccggaggct	4020
gctgttcccg	cggtggaaatc	caagtgggg	accctggaa	ccttgcgttgc	aaagcatatg	4080
tggaaatttca	tcacgggttgc	acaatatttc	cgagggttgc	caactctgtc	tgcaacccat	4140
ggatagatcat	caactgtatgc	attccaccc	tatccatcca	cccccgttac	cacccaaat	4200
acccttcgtt	ttaacatctt	gggggtatgg	gtggccccc	aacttgc	tcccaaggt	4260
gcttcgttgc	tcgttaggc	ccgcacatcc	ggagcgctgt	ttggcagatc	aggcccttgg	4320
aagggttgc	tggatgtatc	ggggatgtt	ggagggatgg	tgccggggcc	gtctgtgtcc	4380
ttaaaggatca	tgacggggca	gtatgcgtcc	accggggac	tgtttaacat	actctgtgt	4440
atcccttc	ctggccgtt	agtcgttgc	gtgtgtgtgc	cagcgatact	ggttcggccac	4500
gtggggccat	gggggggggc	tgtcgtatgg	atgaacccgg	tgatagtcgt	cgcttcggcc	4560
ggtaacccacg	tcccccac	gcatactatgt	cttgcgtatgg	acgctgtcc	acgtgttact	4620
cagatcttc	ctatgttac	catactatcg	ctgctgttgc	ggcttcacca	gtggatcaac	4680
gaggactgt	ccacggccat	ctccgggtc	tggctaaag	atgtttggta	ttggatatgc	4740
acgggtttgtt	ctgatattca	gacctgggtc	cagtccaa	tcctgcgtcg	attgcgggaa	4800
gtcccccttc	tctatgttca	acgtgggtac	aaagggttac	ggccggggcc	cgccatcatgc	4860
caaaacccat	ggccatgttgc	agacacatgc	acccggat	tggaaaacccg	ttccatgttgc	4920
atctgtgggc	cttaggcgtt	tagtaacac	tggcatggaa	cattcccat	taacgcgttac	4980
accacggcc	cctgcaccc	ctcccccgg	ccaaat	ctagggcgt	gtggccgggt	5040
gctgttgc	agtagctgtt	ggttacccgg	gttggggatt	tccaaactgt	gacggggat	5100
accactatgt	acgtatgttgc	cccggttgc	ccgttgcacaa	ccttccatgc	cacaaatgttgc	5160
gatgggggttgc	gggttgcacac	gtatgttgc	ccgttgcacaa	ccgttccatgc	ggagggatgt	5220
acattctgg	tcggggccat	tcaatacc	gttgggttac	agtcctccat	cgagcccgaa	5280
ccggacatgt	cagtgttgc	tttcatgttc	accggatcc	ccacatattac	ggccggagac	

gctaaagcgtta	gggtggccagg	ggatctccc	ccctccttgg	ccagctcatc	actatccatcg	5340
ctgttcgtccgc	cttcctttggaa	ggcaacatgc	actaccgcgc	atgactcccc	ggacgtgtac	5400
cttcatcgggc	caaaaccttct	tggtggccag	gataggatggc	ggaaacatcc	ccgcgtggag	5460
tcgaaaaataa	aggtagtaat	tttggatgtt	ttcagagccgc	tccaagegga	ggaggatgtag	5520
aggaaagtat	ccgttccggc	ggagatcttg	ggatgttcc	ggaaatattc	tcgacggatc	5580
ccccatatacg	cacggccggg	tatcacccct	ccactgttgg	actgtctggaa	ggacggccgt	5640
tcgttcctcc	tcgttccggg	tttggatgtt	tttggatgtt	ccaaaggcccc	tccgtatccca	5700
ccttcacggaa	cgatgttccaa	cggtgttccaa	ttcggatctgg	cgctgttcc	tccgttggcg	5760
gagctcgccaa	ggaaaggagac	gggtgttccgg	ttcggatctgg	cgctgttcc	cgctgttcc	5820
aggccatcttc	caaaagaccc	cggtgttccgg	ttcggatctgg	cgctgttcc	cgctgttcc	5880
tcctccatgc	ctggggccgg	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	5940
tctacatggaa	gggggggggg	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6000
acaggccggcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6060
agaacaatctt	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6120
ctggggggaaa	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6180
tggtgtcaagg	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6240
gaagccctgtt	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6300
gacgtccgggg	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6360
ctggaaagaca	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6420
gtccaaaccagg	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6480
gttctgtgtt	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6540
atggcccttc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6600
gctctggaaag	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6660
acgttgttttt	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6720
cccgaaagoca	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6780
acttaatcttca	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6840
accaggcttgc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6900
aaggccatggaa	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	6960
ggggggggggcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7020
tctggccccc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7080
tccttcataatg	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7140
gacccaccaaa	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7200
tctctggcttag	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7260
actctatcttc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7320
atctatccgggg	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7380
catggccctta	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7440
tcatgttccca	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7500
gtccggcgta	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7560
aacttggccagg	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7620
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7680
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7740
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7800
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7860
tagtgttggaa	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7920
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	7980
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8040
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8100
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8160
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8220
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8280
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8340
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8400
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8460
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8520
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8580
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8640
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8700
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8760
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8820
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8880
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	8940
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	9000
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	9060
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	9120
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	9180
tttggatgttcc	tttggatgtt	tttggatgtt	tttggatgtt	gtatgttcc	cggtgttcc	9240

aacgggttgg catggattgtt	aggccgcgcgc	stataccttg	tctgccc	cgggttgggt	9300
cgcgggtcgat	ggagccgggg	cacctcgacc	tgaatggaa	ccggggcac	9360
gatccacca	ttccaagaatt	ggggacatc	aatttcgtgg	gagaactgtg	9420
ccaaaccttgc	gcaaaaaatc	tccatcgct	ccggccatcc	cagcggccgc	9480
tctcgggcag	ctgtgggtcc	tggccacggg	tgecatgtat	cgtgttctg	9540
ccccggctgg	ctggccgggt	tggcttaactg	gttagcaga	tgaatccacg	9600
gaacgttgg	cgactctgt	tgcacaaatc	ctggacactt	agcaacaaca	9660
tccgttccg	tgtttcgtaa	agtctggaaa	ccggggaaatc	atacgcgac	9720
tcccgatctg	catcgccagg	tgtctgtgc	taccctgtgg	accatattatgt	9780
cgaaggcgt	gtatggaccc	tgtgttattt	ttctctgttc	ccggccatcc	9840
gttgttacc	ctcacaaatc	tccatgttacc	gggcgttcc	atcatcagta	9900
tgatccatct	cttcgttctt	atccgttata	ttacccttac	acccgtatcg	9960
acggaggcat	caatgttgc	aacaggaaaa	aaccgcctt	aacatggccc	10020
aaggccagat	ttaacatgg	tggggaaaat	caacgatgt	gacgggtatc	10080
catctgtgaa	tcgttccatc	accacgcgtg	ttaggttttc	ccgcgttgc	10140
cggtgatgc	gggtggaaacc	tctgcacatc	ggatccccc	ccatgttgc	10200
gtaaaggcgat	ggccggggatc	gacaaggccg	tcaggccgg	tcagggggtt	10260
tcggggccat	gccccatggcc	agtcacgtat	cgatagcgga	gtgttatact	10320
ggccgcattcg	agcagatgt	atccgtatgc	cacccatgtc	gtgggttata	10380
tgccgttgg	ggaaatatacc	catcgccgg	tccctgtttt	ctcgcttac	10440
cgctcggtcg	tteggctcg	gggggggtt	tcagtcgtact	caaaaggccgt	10500
tccacatgg	cggggggatc	egccggaaa	aaatgttgc	aatacgttca	10560
aggaaacccgt	aaaaggccgc	gttgcgttgc	tttttcata	ggctccccc	10620
catcacaaaat	atccgtatcc	aaatcgagg	tggccggaaa	ccatcgatc	10680
cagggttgc	cccccgttgg	ctccctctgt	cgcttcctcg	ttccgaccc	10740
ggatctgtt	ccgcctttt	cccttcggaa	agcgttggcc	tttctcatcg	10800
aggtatcttc	gttgcgttgc	gttgcgttgc	tccaaatgtt	cttcacgttgc	10860
gttcacccgc	accgcgttgc	cttattccgt	aactatgtt	tttgcgttca	10920
caacgttatt	ccgcacatgg	agcggccact	ggtaacagg	tttgcgttgc	10980
ggcggttgc	cagatgtttt	aaatgttgc	ttaactactt	gatcactatc	11040
ttttgtatct	ggcgttgcgt	aaatcgatc	acccatggaa	aaaggttgc	11100
tcccgcaaac	aaaccacccg	tgttagccgt	gggttttttgc	tttgcgttgc	11160
cgccggaaaa	aaggatcttc	agaatccatc	tgtatcttgc	ccacgggttc	11220
tggaaacaaa	actcaacgtt	agggatttttgc	tcatgagat	tatcaaaaatc	11280
tagatccccc	tcttagataat	acgacttact	ata	cccgatcc	11313

<210> 6
<211> 11313
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 6	60					
ggccgcggcc	gattggggcc	gacactccac	catatagtcac	tcccgttgc	ggaactactg	60
tcttcacggca	gaaaaggctt	agccatggcg	tttagtatgg	tgtctgtcag	cctccaggac	120
ccccctcc	ggggagccca	taggttgc	ccggggccgt	gagtagatcc	ggatttgcac	180
gacgacccgg	tccttttgc	gtacaaatcc	cttcaatgc	ggagattttg	ggctgcccc	240
gcgagactgc	tagccggta	gtgttgggtc	ggggaaaggcc	ttgtgttact	ggctgtatgg	300
gtgttgcgaa	gtgtcccggg	agggttctgc	gacgttgcac	catggacacg	aatccatcaa	360
ctcaaaaggaa	aacccaaatgg	ccggccatgt	ttaacaatgg	ttgatggcc	ggagggttcc	420
cggcgcgtt	gggtggaggg	cttattccgt	atgacttgc	acaacagaca	atcggttgc	480
ctgtatccgc	ctgtttccgg	ctgttcaggc	aggggggccc	ggtttttttgc	ttcaagaccc	540
acctgttccgg	ttccctgtat	gaactgcagg	acggggccgc	ggggctatcg	tggctggcc	600
cgacggggct	tccttgcggc	gtgtgttgc	acgttgcgtac	tgaagggga	agggacttgc	660
tgcttattgg	ccggatgtcc	ggggaggatc	tccctgtatc	ttccctgttgc	tttgcgttgc	720
aagtatccat	catggctgtat	gcaatggcc	gggtgtatcc	tttgcgttgc	tttgcgttgc	780
cattcgacca	ccaaatggaaa	catcgatcg	acggacatcg	tactcgatgc	aaaggccgttgc	840
tttgtatctt	ggatgtatcc	gacggatgg	atccatgggg	tttgcgttgc	tttgcgttgc	900
ccagggttca	ggccgttccat	ccggggccgc	aggatccatcg	tttgcgttgc	tttgcgttgc	960
gcttgcggaa	tatcatgttgc	ggaaatggcc	gggttttttgc	tttgcgttgc	tttgcgttgc	1020
tgggtgttgc	ggaccgtat	caggatcatag	tttgcgttgc	tttgcgttgc	tttgcgttgc	1080
ttggccggcga	atggggctgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	1140

agccgcatacg	cttctatcgc	ctttttgaacg	agtttttttg	agtttaaca	gaccacaa	cgc	1200
gtttccctct	agggggatca	atcccccccc	tcccttcoc	cccccccta	cgttactggc		1260
cgaaggcgc	ttgttaaa	ggccgtgtgg	tttgcgtt	ttgttat	tttccat		1320
ccgtctttt	gcaatgttag	ggccggggaa	cttggccctg	tcttcttgc	gacat		1380
aggggcttt	ccctctcgc	caaaggat	caaggctgt	tgaatgtcg	gaaggaa	gca	1440
gttccctgg	aaggcttttg	aaagaaaa	acgtctgt	cgacccttt	caggcagg	ggg	1500
aaccccccac	ctggcacac	gtgccttc	ggccaaa	cacgttata	agat	actacc	1560
gcaaggogg	cacaaccca	gtgcccagt	ttgagttgg	tagttgtgg	aaggat	caaa	1620
tggctctct	caacgtt	caacaagg	cttggagg	ccccaa	atcccattt	gt	1680
atggatct	atctggggcc	tcggcaca	tgctttacat	gtgttt	gaggtaaa	aa	1740
aacgtttag	ccccccaa	tcggggac	tggttttct	tttggaaa	cgat	atacc	1800
atggccctt	ttacggcct	cttccaa	acggcagg	tactttgtt	cat	catact	1860
agcctcac	ggccggac	gaaaggc	gaggggg	tccaa	gttgc	cttccac	1920
acacaatct	tccggcag	ctggcgt	ggcgttgg	ggactgtt	tcat	gttgc	1980
ggctcaaaa	cccttgcgg	ccccaa	ccatcac	aatgt	caatgt	ggac	2040
caggatctcg	tcggcgtc	agggcccc	gggggg	cttgcac	atgcac	cttc	2100
ggcagcgcgg	acctttact	gttccagg	ttccgcgt	tcattcc	gcccgg	ggg	2160
ggcagacaca	gggggagct	actctcc	aggccc	cttactt	gggtt	cttgc	2220
ggcgcgttac	ttgttgtcc	tcggggc	gtgttgg	tcc	tttgc	gggttgc	2280
acccgggggg	ttccggaa	gggttgg	gttccgtt	actat	tttgc	gggttgc	2340
gggtttccgg	tcttcacgg	caactgc	ttccggcc	tacccgc	atcc	agggttgc	2400
gccccatctac	acggcccc	ttgttgc	aaagacta	agg	ttcc	atgc	2460
gccccaa	ataaatgt	ttgttgc	ttgttgc	ccgc	ccat	tttgc	2520
gctgtatgtl	cttacggc	ttgttgc	ttgttgc	tttgc	tttgc	tttgc	2580
accacgggtt	ccccatcac	gtacttcc	tatggca	tttttgc	tttgc	tttgc	2640
tctgggggg	cctatgc	cataatat	gttgc	acttca	acttca	acttca	2700
atcttgcgtt	ttccggac	cttgcac	tttgcac	tttgc	tttgc	tttgc	2760
ctcccccac	cttgcgtt	ggatgtt	tttgcgtt	tttgc	tttgc	tttgc	2820
gtctgttca	gtactgtt	aaatcc	tttgcgtt	tttgc	tttgc	tttgc	2880
aaggggggg	ggccatctt	tttgcgtt	tccaa	aat	tttgc	tttgc	2940
aagtgttgc	gttccgt	caatgtt	gttgc	tttgc	tttgc	tttgc	3000
ataccaactt	ggccggac	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3060
ggcgattttc	actcaatgt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3120
ctggacccg	catttcacat	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3180
cagcgggg	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3240
gaacggccct	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3300
gcttgggtac	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3360
ccagggttgc	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3420
accacatag	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3480
ctggtagat	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3540
caaatagtt	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3600
tataggttgc	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3660
atggcatgt	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3720
gtcttagat	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3780
atcatctgtt	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3840
gtatgtat	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3900
gaacatttca	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	3960
gctgttcccg	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4020
ttaaatttca	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4080
gctgtatgt	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4140
acccttctt	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4200
gtttctgtt	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4260
aagggttgc	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4320
ttaaatttca	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4380
atcccttcc	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4440
gttggccctt	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4500
ggttaaacac	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4560
cagatctctt	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4620
gaggacttgc	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4680
acgggttgc	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4740
gttcccttct	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4800
caaaccat	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4860
atctgttgc	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4920
accacgggg	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	4980
gtctgttgc	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	5040
accacttgaca	tttgcgtt	tttgcgtt	tttgcgtt	tttgc	tttgc	tttgc	5100

atgatttc togttccgg	ggcatatggg	atgcggcgt	tgcaggccat	gctgtcoagg	9120
caggtagatg acgaccatca	gggacaggtt	caaggatcg	tcgggggtct	taccagccata	9180
acttcgtatc ctggacgtt	gatcgtca	cggtttatcg	ccggctcgcc	gagoacatgg	9240
aacgggttgg catggatgt	aggccccggc	ctataccctt	tctggctccc	cgcggtcggt	9300
cgcggtcgat ggagccggc	cacctcgacc	tgaatggaa	ccggggggac	ctcgctaa	9360
gattcacac tccaagaatt	ggagccaaat	aatttcgtcg	gagaactgt	aatgcgc当地	9420
ccaacccctt gcagacata	tccatcggt	ccggccatcc	cagcagccgc	acggcggccg	9480
tcttggccgg	cggttgggtcc	tgcccacg	tgccatgt	ctgtgtcc	9540
cccggttgg	tcgggggtt	tgccattact	gttagcaga	tgaatcccg	9600
gaacgttgg	cgactgtcg	tgc当地aa	ctggccatcg	agaacaaca	9660
tcgggttccg	tggttccgg	agtcgtggaa	ccggggatgc	ccggccctgc	9720
tcggatctc catcgccgg	tgctgtcg	tacccctgtt	aacacttac	tctgtatata	9780
cgaaaggctg	gtatgtaccc	ttctdtgtt	ccgcgc当地	cataccggcc	9840
gttggtttac	cttacaa	tc当地taacc	gggcatgtt	accgc当地	9900
tgagcatcc	cttcgttgc	atccgttata	ttacatcgat	ttacatcgat	9960
acggggatc	caatgttcc	aaacggggat	aatatggccc	gctttatcg	10020
aaggcagaca	ttaaacgttcc	ttggaaaact	caacggatgt	gaggccgtat	10080
catctgtgaa	tcgatccac	accacgtgt	tgagtttac	ogcagctgt	10140
cggtgtatgc	ggtagaa	ttctgttcc	ccgtacatcc	gagacgttca	10200
gttagccggat	ggccggggat	gacaaggccg	ttaggccgg	tc当地gggt	10260
tcggggccg	gccatggcc	agtcaatgt	cgatggccg	gtgtatact	10320
ggccatcatcg	agcagatgt	atgcggatgt	cacccatgtc	gttggtaat	10380
tcggtaa	ggaaaatccat	catcgccgg	tcttccgtt	ctctgttcc	10440
cgctcgatcg	ttccgttgc	ggggggatgt	tc当地atgt	caaaggccgt	10500
tc当地acat	cggggatata	cgccggaaat	aatatgttgc	ataacgttca	10560
aggaaacgtt	aaaaggccgc	tttgcgtgg	ttttccata	ggctccgccc	10620
catcacaat	atgcgttcc	aatgc当地gg	ttggggaa	ccctgtggag	10680
cagggttcc	ccccctggaa	ttccctgttgc	cgcttcttgc	ttccgaccc	10740
ggatctatgt	ccggccctt	cccttccgg	agcgtggcc	tttcttcat	10800
aggtatccat	gtccgttgc	ttggatgttgc	tccaaatgtt	gtctgttgc	10860
gttcagccg	acgcgttgc	tttacccgtt	aactatgttgc	tttgc当地aa	10920
caacgttat	cgccgtatcg	agcggccat	tttgc当地gg	tttgc当地at	10980
ggccgtgttgc	caatgttcc	aaatgttgc	tttgc当地at	tttgc当地at	11040
tttgc当地at	tgctgttgc	aaatgttgc	tttgc当地at	tttgc当地at	11100
tc当地ggaaac	aaacccacgc	ttgttgc当地	gggtttttt	tttgc当地aa	11160
cgccggaaaa	aaaggatctt	aaagatccat	tttgc当地at	ctacggggcc	11220
tggaaacggaa	acttcacgtt	agggttttgc	gtc当地at	tgacgttgc	11280
tagatccccc	tctagataat	acgacttacat	ata	gatcttcc	11313

<210> 7
<211> 11131
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 7 /					
gcggcccccc	gatgggggg	gacactccac	catagatcac	tccctgttgc	ggaactactg
tcttcacgc	gaaaggcgtt	agccatgggg	tttagatgt	tgctgttgc	cctccaggac
ccccccctcc	ggggagggca	tagtggatc	cgggacccgt	gagtagaccc	gaatttgc当地
gacgacccgg	tcccttttgc	gtacaa	cttacatgtt	ggagattttg	gctgtcccc
gcgagatcg	tagcgtatgt	ttgttgc当地	ggggaaaggcc	tttgc当地at	ggctgtatagg
gtgttgc	gtggccccc	agggttgc当地	gaccgttgc	catgacgtat	aatcttac
ctcaaa	aaaccaaaagg	cgccatgt	ttgaaaca	ttgatgttgc	gcagggttgc
cgccgttgc	ggggggagg	tttgc当地at	tttgc当地at	tttgc当地at	480
ctgtatgc	cggttccgg	ctgttgc当地	tttgc当地at	tttgc当地at	540
acccgttgc	ttgttgc当地	tttgc当地at	tttgc当地at	tttgc当地at	600
cgacgggggt	tccttgc	tttgc当地at	tttgc当地at	tttgc当地at	660
tgcttgc当地	ttgttgc	tttgc当地at	tttgc当地at	tttgc当地at	720
aagtatccat	catgttgc	tttgc当地at	tttgc当地at	tttgc当地at	780
catttgc	ccaaacgtt	tttgc当地at	tttgc当地at	tttgc当地at	840
ttgttgc	ttgttgc当地	tttgc当地at	tttgc当地at	tttgc当地at	900
ccagggttcc	gggttgc当地	tttgc当地at	tttgc当地at	tttgc当地at	960

gcttgcggaa tatcatgggt gaaaatggcc	gctttttcgg attcatcgac	tgtggccggc	1020
tgggtgtgg ggaccgtat caggacatacg	cgttgctac cgctgatatt	gctgaagagc	1080
ttggggcggc atggggtagc	cggttcctcg ctgtttcaegg	cccgatttcg	1140
agccatcgc cttttatcgc	cttcttgcacg agttttctcg	agttaaaaaca	1200
gtttccctct	agcggatca atccgcggcc	tccctccccc	1260
cgaaaggcgc	tggataaaaggc	ccccccctaa	1320
cgctttttg	cggtgtggc	tgttattttc	1380
agggttctt	ccctcttcgc	caccatattg	1440
gttccttgg	aaaccttgg	cgatcttcgt	1500
aacccccac	ctggcgacag	cgaccctttg	1560
gcaaggcggg	cacaacccca	caggcagogg	1620
ttgttctct	cttgcacatgt	atgttgttgg	1680
atggatcttg	atccggggcc	cccaagggtt	1740
aaegtatagg	tcgggtcaca	tgcttttctt	1800
atggcgccct	tcgggggg	tggtttttct	1860
agccctcacg	ggccggacag	catatcact	1920
acacaatctt	tcttggggac	cgccgttgc	1980
ggctcaaga	cccttgcgg	ggccgttgc	2040
caggacatcg	tcgggtcaca	aaatgttacac	2100
ggcagcggc	acccatctt	cgatccggat	2160
ggcgcacacg	ggggggact	tcatttcgg	2220
ggcggttcc	tgtcttgc	ggggccgttt	2280
accggagggg	tttgcgggg	ctacttggaa	2340
cggttcccg	ttttcgggg	ttttttggc	2400
gccccatctt	acccatctt	ttttttggc	2460
gccaagggtt	ataagggtgt	ttttttggc	2520
ggttatatgt	ctaaggcaca	ttttttggc	2580
accacggggg	ccccccatctt	ttttttggc	2640
tctgggggg	cttcatatgt	ttttttggc	2700
atccgttggc	tcggcggat	ttttttggc	2760
ctcgccaccc	cttgcgttcc	ttttttggc	2820
gctctgttcc	gactgttgg	ttttttggc	2880
aagggggggg	ggccatctt	ttttttggc	2940
aaggttctcg	gcttcggact	ttttttggc	3000
ataccacta	cgccggagct	ttttttggc	3060
ggcgattttc	acttcgtat	ttttttggc	3120
ctggccatcc	cttgcacat	ttttttggc	3180
caggccggc	tttgcgggg	ttttttggc	3240
gaacggccct	cgggatctt	ttttttggc	3300
gtttgttgc	agtttcgg	ttttttggc	3360
ccagggttgc	cggttgc	ttttttggc	3420
accacccatag	aecccattt	ttttttggc	3480
ctgttgcacat	ttttttggc	ttttttggc	3540
caaataatgtt	ttttttggc	ttttttggc	3600
tataggttgg	ggccgttca	ttttttggc	3660
atggcatcg	tgttcgttgc	ttttttggc	3720
gttcttagcg	cttgcgttgc	ttttttggc	3780
atccatctgt	ccggaaaggc	ttttttggc	3840
gatgatgttgg	aaaggatgtgc	ttttttggc	3900
gaacatccat	aaacagggtt	ttttttggc	3960
gctgtccccc	cggttggat	ttttttggc	4020
tggatatttc	tcacgggtt	ttttttggc	4080
gatgtatgt	cactgttgg	ttttttggc	4140
accctcttgc	tttacatctt	ttttttggc	4200
gtttctgtt	ccggatctt	ttttttggc	4260
aagggttctt	ttttttggc	ttttttggc	4320
ttaaggatca	ttttttggc	ttttttggc	4380
atcccttcc	cttgcggcc	ttttttggc	4440
gttggcccaat	gggggggg	ttttttggc	4500
ggttacacg	tttcccccc	ttttttggc	4560
cagatcttc	ctatcttc	ttttttggc	4620
gaggactgt	ccacgcac	ttttttggc	4680
acgggttgg	ctgatcttc	ttttttggc	4740
gtttcccttc	ttttatgttc	ttttttggc	4800
caaaccacat	gccccatgg	ttttttggc	4860
gttggccggc	tttggatctt	ttttttggc	4920

atgatcgcc	tgtcgctgc	ggtattcgg	atcttcacg	cctctgcgtca	agccttcgtc	8940
actggcccg	ccaccaaacy	tttcggcgag	aaggcggca	ttatccgggg	catggccggc	9000
gacgcgtgg	gtaeactt	gttggggttc	gogacgggg	gctggatgg	cttccccat	9060
atgattttc	tcgttccgg	cggtatccc	ttggccgg	tgccggatgt	gtgttccagg	9120
caggtagat	acgaccata	gggacaggct	caaggatgc	tcggggctct	taccagctt	9180
acttcgcga	ctggacgct	gtatgc	ggatgtat	ccgcctcg	gaggcacatgg	9240
aacgggttg	catgtatgt	aggcgccgc	ctatactt	tctgcctcc	cgttgtgtt	9300
cgcggtcat	ggagccggc	cacctcgacc	tgaatggaa	ccggggcac	ctgcgttaacg	9360
gattccac	tccaagaatt	ggagccaa	aatttcttgc	gagaactgt	aatgcggcaaa	9420
ccaacccct	gcaaaaaata	tccatcgct	ccggccat	cagcagccgc	acgcggggca	9480
tctggcccg	cttggggcc	tggcccccgg	tgccgtatgt	ctgttctgt	tcgttgcgg	9540
cccggttag	ctggcgccgg	tgcctactt	gttagcaga	taatacacc	atacgccgac	9600
gaacgtgaag	cgatgtgtc	tgaaaaacgt	ctgcgacgt	agcaacaaca	tgaatggct	9660
tcgggttcc	tgtttcgaa	ctggggatgg	ccggccctgg	accattatgt	9720	
tccggatct	catcgccgg	tgcgttgc	taccccttgc	aacacttac	tctgttattt	9780
cgaaggcgt	gatgtggacc	tgatgttt	ttctctgttc	ccggccatc	cataccggcca	9840
gttgttacc	ctcaacaacgt	tccgttaacc	gggcgttgc	atcatcgat	acccgtatcc	9900
tgagcatct	ctctcggtt	atcggtatca	ttacccttac	gaacagaat	tcggcccttac	9960
acggggat	caagtgcac	aaacggaaaa	aaccggcc	aatatggcc	gctttacttgc	10020
aaggccagaa	ttaacgttcc	tggggaaact	caacgacgt	gaecggatgt	aacaggcaga	10080
catctgtaa	tgcgttcaag	accacgtca	tgatgtttac	cgccacgtgc	tcggccgtt	10140
cgggtgatgc	gggtggaaa	tgcgttgc	cgatgtcc	gagacggta	cagcttgcgt	10200
gttaacgggt	ttggggggcc	gacaacggcc	tcaggggccg	tcaggggtt	ttgggggggt	10260
tcggggggcc	ccatgcac	agtcgtatgt	cgatagcgg	gtgttatact	gttttactat	10320
gccccatcg	agcgtatgt	actggatgt	cattatgc	gtgttgaat	acccgcacaga	10380
tgctgtatgg	aaaaataccg	catcaggcc	tttccgtt	cctcgatc	tgactcggt	10440
cgccgttgc	tttgggtgc	gggggggtt	tcaggttcat	aaagggggt	aatacggtt	10500
tccacagaaat	caggggatata	cgccggaa	aacatgtgt	aaaaaggcca	gcaaaaggcc	10560
aggaaacgta	aaaaggccgc	tttgcgtgg	tttttccat	ggctccggcc	cccttgacgag	10620
catcacaaaa	atcgacgtc	aagtccagg	tggccggaa	cgacgaggat	ataaaatgc	10680
cagggttcc	cccttggaa	ctccctgtgt	cgctctctg	ttccgaccc	gcccgttacc	10740
ggatcttgc	ccgcgttcc	cccttgggg	agcgttggcc	tttcttcat	ctcagctgt	10800
aggatcttc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	10860
gttcagcccg	accgtcgcc	tttateccgt	aaatctcg	tttgcgttca	cccggttaaga	10920
cacgacttat	cgccatcgcc	agcggccat	ggtaacagg	tttgcgttca	cccggttaaga	10980
ggccgtgtca	cgaggtttt	ggatgtgtt	cctaactac	gtctactac	aaggacatgt	11040
tttgcgttgc	gggttttttt	gggggggg	atcttccgg	aaaggtttt	tagtttttt	11100
tccggcaaac	aaacccacc	ttgttgggt	gggttttttt	tttgcgttca	gcaaggattac	11160
cgccggggaa	aaggatctca	agaagatct	tttgcgttca	tttgcgttca	tttgcgttca	11220
tggaaacggaa	acttcacgtt	agggatttt	gttgcgttca	tatcaaaaa	gattttcc	11280
tagatccctt	tctagataat	acgactca	ata			11313

<210> 8
<211> 11313
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 8

gccagcccc	gatggggggc	gacactccac	catagatcac	tccctgtga	ggaactactg	60
tcttcacgca	gaaaggctct	agccatcgcc	ttatgtatgt	tgtcggtcg	cctccaggac	120
cccccttccc	ggggagagcc	tagtgcgtt	cggaacccgt	gagatcaccc	gaatttgcac	180
gacgacccgg	tcctttttcg	gtacaaacccg	cttaatgcct	ggagattttg	gcgttcccc	240
gcgagactgt	tagcgatgt	gttgcgttgc	ggggaaaggcc	tttgcgtatgt	gcctgtatagg	300
gtgttgcgtt	gtgtcccccgg	aggctctgtt	gacccgttgcac	catcgacacg	aatcttaaac	360
ctcaaaagaaa	aaaaacccagg	cgccgtatgt	ttgaacaaga	tggatgttgcac	gcagggttcc	420
cgccgttcc	gggggggggg	atcttgcgtt	atgtactggc	aaacacacaa	atcggtgtct	480
ctgtatccgg	ctgttccgg	ctgtcgacgg	agggggccgg	gggttttttt	gtcaagacccg	540
acctgtccgg	ttccctggaa	gaactcgagg	acggaggccgg	gggggtatgt	ttgggtggccca	600
cgacgggggg	ttccctggaa	gtgtgtgtcc	acgtgtgtca	tgaagggggg	aggggacttggc	660
tgtatccgg	cgaaatgtcc	ggggcaggatc	tccgttgcata	tcacccgttgc	cctgcgtatgt	720
aagtatccat	catgggtgtat	gcaatggggc	ggctgtatgt	gttgcgtatgt	gttacccgttgc	780

cattcgacca	ccaaaggaaa	catcgatcg	agcgagaacg	tactcgatgcg	gaagccggtc	840
ttgtcgatca	ggatgatctg	gacgaaaggc	atcagggggt	cgcgcagoc	gaactgttcg	900
ccaggctcaa	ggggcgatcg	ccccggggcg	aggatctcg	cgtgaccat	ggcgatgtct	960
gcttgcgcaa	tatcatcgaa	aaaaatggcc	ggttttcgt	atcatcgac	tgtggccggc	1020
tgggtgtggc	ggacgcgtat	caggacatag	cgttgctac	cgtgatata	gctgaagagc	1080
ttggccggca	atgggtcgac	cgtttccctcg	tgcttttacgg	tatcgccgt	cccgatgtcg	1140
agcgcatacg	cttctatcg	cttcttgcag	agttttctcg	agttttaaaca	gaccacaaacg	1200
gtttccctct	aggggggatcg	atccggccccc	tcccttcgg	ccccccctta	cggtactggc	1260
cgaagcgcgt	tggataaaagg	cgggtgtcg	tttgctata	tgttattttc	caccatattt	1320
ccgttctttg	gcaatgttag	ggccggggaaa	cctggccctg	tcttttgc	gaggattctct	1380
agggtttttt	cccccttcgt	caaaaggatg	caaggctgt	tgatgtcg	gaaggaaaggc	1440
gttctcttgg	aaggttttcg	aaagaaaaac	acgtctgt	cgacccttgc	caggcaggcc	1500
aaccccccac	ctggcgcacag	gtgcctctcg	ggccaaaacg	cacgtgtata	agatacacat	1560
gcaaaaggcg	cacaacccca	gtggccatcg	tggtgtgg	tagttgtgg	aaggatcaaa	1620
tggctcttcc	caacggatcg	caacaaagggg	cttggaggat	ccccaaagggt	accccaattgt	1680
atgggtatcg	atctggggcc	tcgggtgcaca	tgcttttgc	tggtttatgc	gagggtttaaa	1740
aacgtcttag	cccccccgaa	cggggggacg	tggttttct	ttggaaaaaca	cgatataacc	1800
atggcgccct	ttacggccct	ctcccaacag	acgcgaggcc	tacttggct	catcatcaat	1860
agcctcacag	ggccggggacg	gaacggatc	gggggggggg	tccaatgtgt	otccaccgc	1920
acacaatcc	tcttggggac	ctggcgatcg	ggcgctgtt	ggactgtcta	tcatgtggcc	1980
gggtcaaaa	cccttgcggg	ccccaaaggcc	ccaatcaccc	aatgtacac	caatgtggac	2040
caggactctg	tcggctgtca	agggccggcc	ggggcgctgt	ccttgcaccc	atgcacccgtc	2100
ggcagcttgc	acccctttatc	gttgcacggg	catggcgat	tcattccgg	ggccgggggg	2160
ggcgcacac	ggggggggact	acttccccc	aggccctgt	cataacttga	gggtttttctcg	2220
ggcggtccac	tgtctcgcc	ctggggccac	tgdtggca	tcttttggcc	tgccgtgtgc	2280
accccgagggg	tttgcggacg	gggtggactt	gtacccgt	tgatctatgg	aaccactatg	2340
cggtttccgg	tcttcggacg	caactcgcc	ccttcgggg	tcggccggac	atttcagggt	2400
gccccatctac	acggcccttc	tggttgcgg	agagacata	aggttgcggc	tgctgtatcg	2460
gccccagggt	ataaagggtgt	tgtctcgaa	ccgttgcgt	ccgcacccct	agggtttccgg	2520
gcttatatgt	ctaaaggccaa	tgatgtatcg	cctaatacata	gaacgggggtt	aaggaccatc	2580
accacgggtt	ccccccatcac	gtactccacc	tatggcaat	tttttgcgca	ogggttgttc	2640
tctggggccg	cctatgtatcg	ataataatgt	tgatgtatcg	actcaatcg	ctcgaccat	2700
atcttggggc	tcggcgacatc	cttggaccaa	ggggggacgg	tttggggccg	actcgctgt	2760
ctgcgcaccc	cttgcgtcc	gggatgggtc	accgtgcac	atccaaacat	cgaggagggt	2820
gctctgtcc	gcacttggaa	ataaaaatcc	tatggcaaa	ccatccccat	cgagaccatc	2880
aagggggggg	ggccaccttt	tttgcgtca	tccaaagaag	aatgtatgt	gctcggccgc	2940
aagctgtcg	gcctcgacat	caatgtgt	gtatattacc	gggggtttctgt	tgatgtatcg	3000
ataaccaata	ggggggacgg	cattgtgtca	gtacccggac	cttataatgc	gggttttacc	3060
ggcgattttc	actcaatgt	gtactgtatc	acatgtgtca	cccagacat	cgactttcgc	3120
ctggccggcc	catttcacatc	tgatgtatcg	acccgtccac	acgggggggt	gtacatgtgc	3180
caggccggcc	ggggggggat	ttggggccgg	atggggattt	acagggtttgt	gactccaggaa	3240
gaacggccct	ccggggatgtt	cgatccctcg	ttttttttcg	tgatgtatcg	cgccgggtgt	3300
gtttttgtcg	agtcgtccac	cgccggggac	tcgtttaggt	tgccgggttt	cctaatacata	3360
ccagggttgc	ccgtctggca	ggaccatctg	gagttctgg	agacgttctt	tacaggccctc	3420
acccacatag	acggccatcc	tttgcgtcc	actaaacgg	caggagacaa	ttttttttttac	3480
ctggtagcat	accaggatoc	gttgcgtcc	agggttccgg	cttcacccatc	atcggtggac	3540
caaataatgt	ggatgtatcg	acggcttaaa	ttttttttcc	tttttttttt	atcggtggac	3600
tatagggtcg	ggccatgttca	aaaaaggatgt	acttccacac	acccccataac	caaataatcata	3660
atggcgatcg	tgttgcgtcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	3720
gttctatgt	cttcggggcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	3780
atcatctttgt	ccggggaaatcg	ggccatccat	ttttttttcc	tttttttttt	atcggtggac	3840
gatgatgtgg	aaaggatgtcg	cttacaccc	ttttttttcc	tttttttttt	atcggtggac	3900
gaacaatcca	aaacggggcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	3960
gctgtcccg	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4020
tggaaatttca	tcagcggtat	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4080
gcgatagatcg	cactgtatcg	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4140
accctcttgc	ttaacatcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4200
gctttttgttgc	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4260
aagggttgcgt	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4320
ttaaagggtca	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4380
atcccttcgg	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4440
gtggccggcc	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4500
ggttaaacacg	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4560
cagatccatct	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4620
gaggactgt	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4680
acgggtttgt	ttttttttcc	ttttttttcc	ttttttttcc	tttttttttt	atcggtggac	4740

gtcccttct ttcatgtca acgtgggtac aaggggatot ggccgggogga cggccatcatg 4800
 caaaacactt gccoatgtgg agccagacatc accggacatg tgaaaaaaacccg ttccatgagg 4860
 atcggtggc cttagactc tagtaacacg tggatcatgg cattcccccatt taaccgcgtac 4920
 accacggggc cctgcacggc ctccccggcg caaaattttt ccattttttttt taaggggcggt 4980
 getgtcgagg agtacgttga gtttacccggc gtggggggat tccatacttcc gacggggatcg 5040
 accactgaca acgttaaatggc cccgttgatc htccggccgc cggaaatttttcccttccatcg 5100
 gatgggggtc gtttgcacac gtacgttccaa tgatggccac agccatccatcg cgaggaggtc 5160
 acatccatgg tgggggttcaaa tcataatccatc tgatgggttc accgcaccccttccatcg 5220
 cggacgttag cagtgttccat ttccatgtcc acccttcctgg ccagcttccatcg 5280
 gctaaggcgtt ggttgcacgg gggatctcc ttccatgtcc accatccctccatcg 5340
 ctctgtcgcc ttccatgttggaa gggaaatccatc accatccatcg atgactccatcg 5400
 ctatcgaggcc caacaaatccatc ttccatgtcc accatccatcg 5460
 tcagaaaaat aggttagtaat ttggatcatc ttccatgtcc accatccatcg 5520
 agggaaatgtt cccgttccggc gggatccgtt ccacttgttag ccacccattac ggcggagacg 5580
 cccatataatggc caacggggggc ttccatgtcc accatccatcg 5640
 taatgttcctt cggatgttaca cgggttgcaccc ttccatgtcc accatccatcg 5700
 cctccacgggaa ggaaggaggac gtttgcacggc ttccatgtcc accatccatcg 5760
 gagetctcgca acggatccatc cggcgttccatc ttccatgtcc accatccatcg 5820
 acggcttcctc ttccatgtccatc ttccatgtcc accatccatcg 5880
 tcctccatggc ttccatgtccatc ttccatgtcc accatccatcg 5940
 tttatccatggc gcgaggaggc tagtggaccc acatccatcg 6000
 acaggccccc acgtatccatc ttccatgtcc accatccatcg 6060
 acgtatccatc ttccatgtccatc ttccatgtcc accatccatcg 6120
 ctggggcaga gaaaggatccatc ttccatgtcc accatccatcg 6180
 gtgttcatggc agatgtggggc gaaagggttccatc ttccatgtcc accatccatcg 6240
 gaaggctgtt acgttgcaccc cccatccatcg ttccatgtcc accatccatcg 6300
 gagtcgttggc ttccatgtccatc ttccatgtcc accatccatcg 6360
 ctggggagaca ttccatgtccatc ttccatgtcc accatccatcg 6420
 gtccaaacccg gtttgcacggc ttccatgtccatc ttccatgtcc accatccatcg 6480
 gtgtcggtgt atgggggttttttccatcg ttccatgtcc accatccatcg 6540
 atgggggtttt ccacatccatcg ttccatgtcc accatccatcg 6600
 gctctggggaa acggatccatc ttccatgtcc accatccatcg 6660
 acggtcaatgtt cccatccatcg ttccatgtcc accatccatcg 6720
 cccggaaaggc acgtatccatc ttccatgtcc accatccatcg 6780
 actaaatccatc ttccatgtccatc ttccatgtcc accatccatcg 6840
 accaggctcgcc ttccatgtccatc ttccatgtcc accatccatcg 6900
 aatgtccatggg ttccatgtccatc ttccatgtcc accatccatcg 6960
 ctatccatgggg ttccatgtccatc ttccatgtcc accatccatcg 7020
 catggggatcc ttccatgtccatc ttccatgtcc accatccatcg 7080
 tctggccccc ttccatgtccatc ttccatgtcc accatccatcg 7140
 ttctccaaatggc ttccatgtccatc ttccatgtcc accatccatcg 7200
 gaccccaacca ttccatgtccatc ttccatgtcc accatccatcg 7260
 ttctgggttag ttccatgtccatc ttccatgtcc accatccatcg 7320
 actatccatctt ttccatgtccatc ttccatgtcc accatccatcg 7380
 atctatccatggg ttccatgtccatc ttccatgtcc accatccatcg 7440
 catggggatcc ttccatgtccatc ttccatgtcc accatccatcg 7500
 ttccatgttccatc ttccatgtccatc ttccatgtcc accatccatcg 7560
 tcatggggatcc ttccatgtccatc ttccatgtcc accatccatcg 7620
 ttccatgttccatc ttccatgtccatc ttccatgtcc accatccatcg 7680
 ctatccatgggg ttccatgtccatc ttccatgtcc accatccatcg 7740
 ttccatgttccatc ttccatgtccatc ttccatgtcc accatccatcg 7800
 ttccatgttccatc ttccatgtccatc ttccatgtcc accatccatcg 7860
 ttccatgttccatc ttccatgtccatc ttccatgtcc accatccatcg 7920
 tagtgtgaa aggtccgtga ggggttccatcg ttccatgtcc accatccatcg 7980
 agatcaatgtt cccatccatcg ttccatgtcc accatccatcg 8040
 gcatatgtac ttccatgtccatc ttccatgtcc accatccatcg 8100
 agtttacatcc ttccatgtccatc ttccatgtcc accatccatcg 8160
 otccatcgta ttccatgtccatc ttccatgtcc accatccatcg 8220
 gtactgtcgcc gcttccatcg ttccatgtcc accatccatcg 8280
 gtgtgtgtatc cggatgttccatcg ttccatgtcc accatccatcg 8340
 tccacccgttcc ttccatgtccatc ttccatgtcc accatccatcg 8400
 tacgatgtca ttccatgtccatc ttccatgtcc accatccatcg 8460
 ggcacatccaccc ggcggccaccc ttccatgtcc accatccatcg 8520
 gaagatccatggg ttccatgtccatc ttccatgtcc accatccatcg 8580
 ggcacatccatggg ttccatgtccatc ttccatgtcc accatccatcg 8640
 gcaatccatggg acggatccatc ttccatgtcc accatccatcg 8700

ggagagacgt	gaccatgcgt	cttgagacgt	ttaaacccat	tcagtttctt	ccgggtggcg	8760
cggggcatgt	ctatcgtcg	cgcaattatgt	actgttttt	ttatcatgtca	actgttagga	8820
cagggtccgg	cagcgtctgt	ggttttttgt	ggcgaggacc	ggcttcgtgt	gagccggacg	8880
atgatccggcc	tgttgtttgt	ggttttttgt	atcttgcgt	cccttcgtca	agcttcgttc	8940
actgtcccg	ccaccaaact	tttccggcgt	aaggaggcca	ttatccgtcc	catggggccc	9000
gacggcgctt	gtcaacgttt	gttgggggtt	ggcgaggacc	gtctgtatgc	cttccccccat	9060
atgattttcc	tcgtttccgt	cggtatccgg	atggccgtgt	tgccggccat	gtgttccagg	9120
caggtagatg	aogaccatca	ggggacgttt	caaggatccg	tcggggcttc	taccacgtcc	9180
acttccatgt	cttgcgttgt	gtatgttcac	gogattttat	ccgcgtccgg	gagcacatgg	9240
aacgggtttt	catggatgtt	ggggccgtcc	atatacttgg	tctgtctccc	cgccgttggt	9300
cgccgtgtat	ggggccgggt	cacctggacc	tgaaatgggg	ccggggccat	ctccgtcaacg	9360
gattcacccat	tccaaagat	ggggccatca	aatttttgcgg	gagaactgtgt	aatgcgc当地	9420
ccaaaccttgc	gcagaacata	tccatgtgt	cgccatcc	cagcggccgc	acggggccca	9480
tcttggggcgt	cggtttttcc	ttggccacgg	tgccgtatgt	cggttgcgt	tggttggagg	9540
ccgggtttgt	ctggccgggtt	ttggccatgtt	ttggccatgtt	ttggccatgtt	atacggcggc	9600
gaacgtgttt	cgactgtgtt	tgaaaaaaact	tttccgttgt	tttccgttgt	ttttccgttgt	9660
tegggtttcc	tgttttgcgtt	agtctggaaa	tttccgttgt	tttccgttgt	tttccgttgt	9720
tcccgatgtt	catcgccgtt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	9780
cgaaaggcggt	gttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	9840
gtttttatcc	cttccaaatcc	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	9900
tgagcatctt	cttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	9960
acggggcat	caatgttcc	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10020
aaggccatgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10080
catctgtgtt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10140
cgggtgtatgt	gggtttttttt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10200
gttaagggtt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10260
tcggggggcc	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10320
ggccatgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10380
tgccgtatgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10440
cgctcggtgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10500
tccacagaaat	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10560
aggaaaggctt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10620
catcacaaat	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10680
cagggttttc	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10740
ggatatactt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10800
aggtatccat	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10860
gttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10920
cacgacttat	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	10980
ggccgtgtat	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	11040
ttttgtatct	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	11100
tcggccaaac	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	11160
cgccggaaaa	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	11220
tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	11280
tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	11313

<210> 9
<211> 11313
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 9

ggccggcccc	gattggggcc	gacactccac	catagatcac	tccctgtgt	ggaactactg	60
tcttcacggc	ggaaaggctt	agccatggcg	tttagatgt	tgttgtgtcg	cctccaggac	120
ccccctccc	ggggagacgt	tagtgggtgt	cggaaacccgt	gagttgtcc	gaattggccag	180
gacggccgg	ttctttttgtt	gtatccatgtt	cttacatgtt	ggatgttttg	ggctggcccc	240
gcccggatgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	300
gttgtgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	360
ctcaagaaaa	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	420
cgccggccgtt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	480
ctgtatggcc	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	540
acctgttgttgg	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	tttccgttgt	600

cgacggggcgt	tcttgcgca	gttgtgtctg	acgttgtcac	tgaagcggga	agggactggc	660
tgttatggg	cgaatggccg	gggcaggata	tcttgtcatc	tcaccctgtc	cctgcgcaga	720
aagtatccat	catggctgt	gaaatggcc	gggtgtatc	gtttgtatcg	gtctacttgtc	780
cattcgacca	ccaaggcggaa	catcgatc	acgcgacag	tactcgatg	gaagccgttc	840
ttgtcgatca	ggatgtatcg	gacgaagagc	atcaggggct	cgccgcacgc	gaactgttgc	900
ccaggctcaa	ggcgcgcgt	cccgacggcg	aggatctcg	cgtgaccat	ggcgatgtct	960
gcttgcgca	tatcatgttg	aaaaatggcc	getttttctg	attcatgac	tgtgcggggc	1020
tgggtgtgc	ggacggcgtat	caggacata	cgttgtgtac	cctgtatatt	gtctgaagac	1080
ttggcggggc	atgggtgtac	cggttctcg	tgctttacgg	tatcgccgt	cccgattcgc	1140
agcgcatcgc	cttctatcg	cttcttgacg	atgtttctcg	agttttaaaaca	gaccacaacg	1200
gtttccctt	atcgccgtatc	atccggcccc	tctccctccc	ccccccctta	cgttactggc	1260
cgaaggcgt	tggataataagg	ccggcgtgtcg	tttgtctata	tgttattttt	caccatattg	1320
ccgttcttg	gcaatgttgc	ggccggggaaa	cttgcctcg	tcttttgcac	gaggatctct	1380
agggttcttgc	ccccctctcg	caaggatgt	caaggctgt	tgtatgtct	gaaggaaacg	1440
gtttctctgg	aaggottttcg	aaagcaaca	acgtgtgt	cgacccttt	caggcaggg	1500
aaccccccac	ctggggcagac	tggtcttc	ggcccaaaagc	cacgtgtata	agatcacact	1560
gcaaggcggg	cacaacccaa	gttgcacgt	tggtgttgc	tagttgtgg	aagagtcaaa	1620
tggctctct	caagcgatatt	caacaagggg	ctgaaggat	cccagaaggt	acccctatgt	1680
atgggatct	atctggggcc	ttctgtcaca	tgctgtatcg	gtgtttatgc	gaggtaaaa	1740
aacgttctagg	ccccccggaa	cacgggggg	tggtttttct	tggaaaaaca	cgataattac	1800
atggogccata	ttacggcata	ctcccaacag	acggcgaggc	tactttgtgt	catatcgatc	1860
agcctcacag	ggccggggac	caaggacgt	ggggggagg	tccaaatgtt	ctccacccgca	1920
acacaatctt	tcttggggac	ctgcgttca	ggcgtgtgt	ggactgttca	tcatgttgcc	1980
ggctcaaa	ccccctggcc	ccccaaaggc	ccaaatccac	aatgttacac	caatgtggac	2040
caggacatcg	tcggctggca	agggcccccc	ggggcgggt	cettgtacac	atgcacccgt	2100
ggcagctcg	acottttact	gttgcacggg	catggcgat	tcatttcgggt	ggccgggggg	2160
ggcggacacg	ggggggatct	actttccccc	aggccgtgt	octactttga	gggttcttcg	2220
gggggttccat	tgttctgc	cttggggc	gttggggca	tttttcgggc	tggcgtgtgc	2280
acccgggggg	ttggggggc	gggtggactt	tgatccgtcg	agttatcgac	aaccactatg	2340
cgggtccccc	tcttc	caacatgt	ctccggggc	tacccggacac	attccagggt	2400
gccccatctac	acggccctac	tggtagggc	aagagacta	aggtgcggc	tgctatgtca	2460
gccccagggt	ataaagggt	gttctgtac	ccgttgcgt	ccocoaccc	agggttccgg	2520
gctgtatgt	ctaaaggcaca	tgttgcacg	cctaatacata	gaaacgggggt	aaggaccatc	2580
accacgggt	ccccatccat	gtacttcacc	tatggcaat	tttttgcgca	cggttgttgc	2640
tctgggggg	cctatgcat	ataataatgt	gtatgtgtc	actcaactga	ctcgaccat	2700
atctggggcc	tcggcacat	cttggacca	ggggggacgg	ctggagcgc	actctgtgt	2760
ctggccaccc	ctacgtcc	ggggatgttc	atccgtgcac	atccaaacat	cgaggagggt	2820
gcttgc	gactgttca	aatccccctt	tatggcaata	ccatccatcc	cgagaccatc	2880
aaggggggga	ggcacccat	ttttgcata	tccaagaaga	aatgttgcgt	gtctccggcg	2940
aagtgcgtc	gtctgtgt	caatgtgt	gtatgttca	ggggcttgc	tgatattcg	3000
ataccaaatc	ggggggatgt	cattgtgt	gaaacggacg	cttcaatgt	ggggtttacc	3060
ggcgatcttc	actactgt	cgactgtat	atctgtgt	ccccacatgt	cgactttcgc	3120
ctggacccccc	cotttacccat	tgagggacgg	acccgtgcac	aaaggccgt	gtcaacgtcg	3180
cagggggggg	ggcggactgg	tagggggagg	atggggat	acagggtttt	gactccagg	3240
gaacatgttgc	ccggggatgt	cgatccatcg	ttttgttgcg	atgtgtatgc	ccgggggtgt	3300
gctttgtatcg	agcttcacgg	ccggcggatcc	tcaggtaggt	tgcgggttca	cctaatacata	3360
ccagggttgc	cgtctgc	ggacccatctg	gtttgttgg	agaggttgc	tacaggcc	3420
acccacatag	acggccat	ttttgtcccg	actaaggc	caggagacaa	cttccccatc	3480
ctggtagat	accaggatad	gggttgcgc	agggttc	cttccaccc	atctgtggac	3540
caaatgttgc	atgtttctat	acggcttaa	cttacgttc	acggggccaa	ggccctgtgt	3600
tatagggttgc	ggccgttca	aaacggatgt	atccacac	acccccataac	caaatacatc	3660
atggcgtca	tgttgcgt	cttggggatc	gtcacggac	ccttgggtgt	ggtagggcga	3720
gtcttagcag	cttggcc	gttgcgt	aaacaggac	gggtgtat	tgtggggagg	3780
atctatgttgc	ccggggatgt	ggccatcat	ccggacagg	aagtcttta	ccggggatgc	3840
gatggatgttgc	aaggatgtgc	ctcacatcg	ttttatcg	aaacgggtat	gcagotcgc	3900
gaacaatccat	aaacggatgt	aaatccgggt	cttccaaacat	ccaccaacgt	agccggggat	3960
gctgtcccg	tgggtgttca	caagtgggg	accctcgaa	ccttctggc	gaagcatatg	4020
tggaaatgttcc	tcacggatgt	aaacatattt	ggggatgtgt	ccactctgtc	tggcaacccc	4080
gctgtatgt	caactgtatgt	atccacatcg	tatccatcca	ggccctcgac	caccccaatc	4140
accctctgttgc	ttaacatccat	ggggggatgt	ttttgttgc	ttttgttgc	ttttgttgc	4200
gctttctgttgc	tctgtatgt	ccggatcat	ggggatgt	tttgcgtat	aggcccttgg	4260
aagggtgttgc	tggatgtatgt	ggggatgtat	gggggggggg	ttggcggggc	gtctgtggc	4320
ttaatgttgc	tgacggggatgt	ttttgttgc	ccggggatgt	ttttgttgc	ttttgttgc	4380
atctcttcgttgc	ctggccgttgc	atgttcgtcc	acccggatgt	ttttgttgc	ttttgttgc	4440
gtggggccat	ggggggggatgt	ttttgttgc	ttttgttgc	ttttgttgc	ttttgttgc	4500
gtggggccat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	4560

cagatccctc	ctagtttac	catcaactcg	ctgtgaaga	ggtttaccca	gtggatcaac	4620
gaggactgc	ccacgcattg	ctccggctcg	tggctaagag	atgtttggga	ttggatatgc	4680
acgggttgc	tcgttccaa	gacggctgc	catggcaac	tcgtccggcg	atgtccggga	4740
gtccccttt	tctcatgtca	acgtgggtac	aaggaggatc	ggccggggga	cggcatatcg	4800
caaaaccac	gcccattgtt	agcacagatc	accggacatg	tggatggaa	cattecccat	4860
atcggtgggc	cttaggactcg	tagtaaacacg	ttccatgtca	taacgcgtac	4920	
accacqgggc	cctgcacgca	ctccgggggg	ccaaattatc	ctaggggctg	gtgggggggt	4980
gctgctgagg	agtacgttgc	gtttagccgg	gtggggggat	tccactaegt	gacggcgtatg	5040
accactgaca	acgttaaatg	cccggtcag	gttccggccc	cggaaatttc	cacagaatgt	5100
gatgggggtt	gggtgcacat	gtacgttca	gogtgcaaa	cccttcatacg	ggaggaggtc	5160
acattctgg	tcgggttca	ttacaatctg	ttggggatcc	agtcctccat	cgagccggaa	5220
ccggacgtcg	cagtgtcata	ttccatgtca	accggaccc	cccaacattac	ggcgagacg	5280
gctaagcgt	ggctggcccg	gggatctccc	cccttccttg	ccagatcata	agctatccag	5340
ctgtctgcgc	cttccttgc	ggcaacatgc	actacccgt	atgtatccc	ggacgctgca	5400
ctcatcgagg	ccaaaccttc	ttggccgcgc	gagatggggc	ggaaacatcc	ccgctgtggag	5460
tcagaaaaat	aggtagtaat	ttttggatgt	ttccagccgc	tccaaaggcg	ggaggatgtag	5520
agggaaatgt	ccgttccggc	ggagatctcg	ccggaggttca	ggaaatcc	tcgagcgtatg	5580
cccatatggg	cacggccggc	ttacaaccc	ccatctgttg	atgtctggaa	ggaccggac	5640
taagtccttc	ctagggttaca	ccgggttcca	ttccggccca	ccaaaggccc	tccgatatacc	5700
cctccacccg	ggaaaggggc	gttgcgtctg	tcgaaactca	ccgtgttctc	tgcccttggcg	5760
gagctcgca	caaagaccc	cggcacgtc	gaatctcg	ccgtcgacag	ccggcggcga	5820
acggcccttc	cttgaccgc	cttgcgac	ggcgcacgg	gttcgcacgt	tgatgtcgat	5880
tcctccatgc	cccccccttg	ggggggagcc	ggggatcc	attcctcgat	ccggcttctgg	5940
tctaccatgc	ggcgaggagg	tagtgcgg	ttccgtctgt	gtctcgatgc	ctacacatcg	6000
acaggccgccc	tgatcgcgg	atgcgtcg	gaggaaacca	agtcgtccat	caatgtactg	6060
agcaactctt	tgtcctgtca	ccacaatcc	gtctatgt	caacatctcg	caggcgaacg	6120
ctggggcaga	agaaggatcg	cttgcacaga	ttccgcgt	ttggacggaca	ctaccggcga	6180
gtgtcaagg	agatggagg	gggggggtc	acagttta	ccaaatcttc	atccgttgag	6240
gaagctgtt	agtcgtgc	ccccatctcg	ggccatcta	attttgtta	ttggggcaaa	6300
gacgtccgga	acattatcc	caaggccgt	aaccatcc	gtccgtgtg	gaaggacttgc	6360
ctggaaagaca	ctgagacaca	aattgacacc	accatcatgg	aaaaaaatga	ggttttctgc	6420
gtccaaacc	agaaggccagg	ccccggcca	gttcgcgtt	ttgttattcc	agatgtggg	6480
gttcgttgt	ggcgaaaatg	ggcccttttc	gtatgttc	ccaccccttc	tcaggccgtg	6540
atgggcctt	catacgattt	ccaaatctc	cttgcacagc	gggtcgatgt	cctgtgttat	6600
gccttggaaag	cgaaagaaatg	ccctatggc	tttcgcata	acaccggctg	ttttactcta	6660
acggtactcg	agatgtacat	ccgttggat	gagtcataat	accaatgtt	tgacttgcc	6720
cccgaaatcc	ggaggccat	aaatgtcg	ccagacggc	tttacatcg	ggggcccccttg	6780
actaaatcc	aaaggccaga	ctggccgtt	ccgcgggttgc	ggcgcggagg	tgtacttgac	6840
accagctcg	gtataatcc	ccatgttac	tttgcggcc	ccgcggccgt	tcgactgtcg	6900
aaatccatcg	actgcacgg	gttcgtatgc	ggagacgac	tttgcgttgc	ttgtgaaaatg	6960
ggggggaccc	aaaggaccc	ggggacgtca	ccggccatca	ccggggatct	gactatgtac	7020
tctggcccc	ctggggaccc	ggcccaaaaa	gtaaatcg	ttccgtgtat	aaatcatcg	7080
tctcccaatg	tgtcgttgc	gtacgtcga	ttggccaaaa	gggtgtacta	tctccacccgt	7140
gaccacccat	cccccccttg	gggggtcg	ttggggacag	ctagacacac	tccagtcaat	7200
tcctggatcg	gcaatccat	ccatgtatgc	ccccacccgt	ggcaaaatgt	gtatctgtat	7260
actatttc	tctccatcc	ttatgttc	ccgttccatcg	aaaaaaatcc	agatgttcag	7320
atctacgggg	cctgttactc	cattgcattc	ttccgtatcg	ccatgtatcg	tcaacacgttc	7380
catggccat	ggccattttc	atccatcgat	tacttcgc	ttggatccat	taggtgttat	7440
tcatgcctcc	ggaaacttgg	ccatccggcc	tttcgcgttgc	ggagacatcg	ggcccaaaatg	7500
gtccggccgg	ggctactgtc	ccagggggggg	gggtgtccca	tttgcgttgc	gtatcttc	7560
aactggccat	taaggccaca	gtccaaatct	actccatcc	ccgttccgtc	ccagtgttat	7620
ttatccatcg	ggttcgttgc	ttgttacagc	ggggggacata	tatatccatcg	cctgttcgt	7680
gccggacccc	gttgggtatc	tttgcgttgc	ttccatcgat	ttgttgcgttgc	aggatctat	7740
ctatccccc	accgtatgac	ggggacatca	acacttcagg	ccaaatggcc	atccgtttt	7800
tttccctttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	ttttcccttt	7860
tttttccot	ttttttccot	ttttttccot	ttgttgcgttcc	atcttagcc	tagtcacggc	7920
tagctgtaa	aggctcggt	gtccgttgc	ttccgttgc	gtctgtatcg	gccttcgtc	7980
agatccatgt	cttcgttgc	ccggccatcg	ttccgttgc	gggggtatgc	ccgggtttta	8040
gcatatttgc	gaccatccat	tcatgttgc	ccgttccatcg	ttccgtatcg	ttatgtcggt	8100
agtttatacg	agtttataatg	cttaacgcgt	caggccacgt	gtatgtatcc	taacaatcg	8160
ctcatcgatc	tcctcgccac	tcgttccatcg	gttcgtatgc	gtatgtatcg	ttgtatgtcc	8220
gtactggccg	ggcttcttg	ggatatcgatc	ttccgttgc	gtatgtatcg	tcactatcg	8280
gtgtctgtatcg	cgatcatgtc	tttgcgttgc	ttttatcgat	ccggccgttcc	ccggacatcg	8340
tccgacccgt	ttggccgcgg	ccccatccgt	ttccgttgc	tttgcgttgc	tttgcgttgc	8400
tacgatgtca	ttggccggacac	accgcgttgc	ttggatccatcg	ccggccgttcc	cactatcgac	8460
ggcatccatcg	ggccacacagg	ttccgttgc	ttggatccatcg	ccggccgttcc	caccatcgatgg	8520

gaagatcggg	ctcgccactt	cggytcatg	agcgcttgyt	tggcggtgg	tatgttgcca	8580
ggcccccgtgg	ccgggggact	gttggcgcc	atctccttc	atgcacatt	ccttgcggcc	8640
gccccgtcga	acggctcaaa	cctacta	ggctgtttc	taatgcagg	gtcgccataag	8700
ggagagcgtc	gaccgtatcc	cttagagcc	ttcaaccgg	tcagtcctt	ccgggtggcc	8760
cggggcatgg	ctatgtcgc	cgactttag	actgtcttc	ttatcatgca	actcgtagga	8820
cagggtccgg	cacggctctg	ggtattcgg	atcttgcgg	ccctgcgtt	gagcgccgacg	8880
atgatcgcc	tgtcgcttgc	tttttttttt	ttatggcgg	actcttcgtc	agcttcgtc	8940
actggtccgg	ccacaaggaa	tttcggcgag	aaggaggcca	ttatcgccgg	atcgccggcc	9000
gacggcggtgg	gtacgtttt	gttggcggtt	ggacgcggag	gttggatggc	tttccccatt	9060
atgtatcttc	tcgttcccg	cggtatggg	atggccgggt	tcggggat	gtctgtccagg	9120
caggtagatg	acggaccatc	ggggacgggt	caaggatcg	tcggggctt	taccaggcc	9180
acttcgtatca	ctggggccgt	gatcgatcg	gogatttat	ccggcttggc	gagccatcg	9240
aacgggttgg	catggatgt	aggcgccgg	cttacatgtt	tcgtctcc	cggttgggt	9300
cgcgggtcgat	ggacgggggg	cacctggacc	tgaatggaa	ccggggggaa	ctcgctaacc	9360
gattcaccaat	tccaaaggat	ggggccata	aatttttgtt	ggagaactgt	aatgcgcaaa	9420
ccaaacccttgc	ggagaatccat	tccatcggt	cgccatctt	cagcaggccc	acggggccga	9480
tctcgccggc	egttgggtcc	tggccacggg	ttcgcatgt	tcgtctctg	tcgttggagg	9540
cccggtttagg	ctggcggggt	tgccctactt	gttagcagaa	tgaatccac	atacgcgagc	9600
gaacgtgtgg	cgactgtcg	tgccaaacgt	ttccggccgt	agcaacaaca	tgaatgggt	9660
tcgggtttccg	tgttgcgtt	agtcgttggaa	ccggggatgt	ccggcccttg	accattatgt	9720
tcoggatctg	catcgaggaa	tgctgttgc	taccctgtt	aacacccata	tctgttattaa	9780
cgaaggcgctg	gcattggacc	tgatgttgc	ttctctgtt	ccggccatc	cataccggca	9840
gttgtttatcc	ctccaaacgt	tccaaatgtt	ggggatgtt	atcatcgat	acccgtatcg	9900
tgagcatctt	cttcgttccat	atccgttgc	tttccccc	agacagaaaat	ttttcccttac	9960
acggaggccat	caagtgtacca	aacaggaaaa	aaccggccctt	aatatggccc	gttttatcg	10020
aaggccagaca	ttaacgttcc	tggagaaat	caacggatgt	gacggggatg	aacaggcaga	10080
catctgtggaa	tgtcttcac	accacgttg	tgatgttttc	ccggactgtc	tcggcggtt	10140
cgggtgtggac	gggtggaaaat	tctgcacat	cgacgttcc	gagacgggtt	cagttgttt	10200
gtaaaggcgat	ggccggggaga	gaaacccgg	tcagggggg	tcagggggtt	ttgggggggtt	10260
tccggggcga	gcccataccat	agtcacatgg	cgataggcc	gttataactt	gtttttactat	10320
gccccatgg	agcagatgt	actggagatg	caccatatgc	ggtgttgaat	acccgcacaga	10380
tcggatggaa	gaaaataccy	catcggccg	ttttccgtt	ccctggctac	tgacttcgt	10440
cgctcggtcg	ttccgggtcg	ggcggcggtt	tcagtcact	caaaggccgt	aatacggtt	10500
tccacagaaat	caggggataa	cgcaggaaag	aacatgttgc	caaaaaggcca	gcaaaaaggccc	10560
aggaaacggta	aaaaaggccg	gttgcgttgc	ttttttccat	ggctccggcc	ccctgtacgag	10620
catccacaaaat	atgcggctt	aaatgcagg	ttggccaaacc	ccagcaggat	ataaaagatcc	10680
cagggttttc	ccctcttgc	ttcccttcgt	tcgtcttcgt	ttccggaccctt	ggccgttttt	10740
ggatccctgt	ccggctttt	cccttcggaa	agcgttggcc	tttccatgt	ctcacgttgt	10800
aggatatctca	gttcgggttgc	gttgcgttgc	tccaaatgtt	gtctgttgc	cgaaaccccc	10860
gttcaggccg	accgttgcgc	tttatttcgtt	aactatgtt	tttgcgttca	cccggttgaaga	10920
cacgactttt	cgccacttgc	ggcggccat	gtttaacagg	tttgcgttgc	gagggtatgt	10980
ggcggtgtca	caggtttttt	gaatgttgc	tttttttttt	tttttttttt	tttttttttt	11040
ttttgttatcc	ggcgtcttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	11100
tccggcaaac	aaaccacccgg	ttgttagcggt	tttttttttt	tttgcacatc	tttttttttt	11160
cggagaaaaaa	aaaggatctca	aaaggatctt	tttgcacatc	tttgcacatc	tttttttttt	11220
tggaaacggaaa	acttcgtatc	aaaggatttt	tttgcacatc	tttgcacatc	tttttttttt	11280
tagatccat	tctatgataat	acgactcact	ata			11313

<210> 10
<211> 11313
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 10						
gccagcccc	gatggggggc	gacactccac	catagatcac	tccctgttg	ggaaactactg	60
tcttcacgca	gaaaggcgct	agccatcgcc	ttatgtatgc	tgtcggtcg	ccttcaggcc	120
ccccctccccc	ggggagaccc	tagtggctt	cggaacccgt	gagttccatc	gaatttgcac	180
gacggactgg	tccttgcgtt	gtatccatcc	tttttttttt	tttttttttt	tttttttttt	240
gacggactgg	tgacccatgg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	300
gtgtttcgca	gtgccccggg	aggtctcgta	gaccgtgcac	catggacccg	aatcttaaac	360
ctcaaaaggaa	accacaaagg	cgccggatgt	tttgcacatc	tttgcacatc	tttttttttt	420

cgggcgttgc	ggggaggagg	ctattccgt	atgactggc	acaacagaca	atcggtctgt	480
ctgtgcgg	ctgttcgg	ctgtcacgcg	agggggccc	ggtttttt	gtcaagaccc	540
acctgtccgg	tcgcgttgc	gaactgtgg	acggaggcgc	gggttatgt	tgggtggcca	600
cgagggggt	tccttcgcgc	gttgcgttgc	acgttgtcgc	tgaaggccgg	aggacttgcc	660
tgttattggg	cgaaatgcgc	gggcaggatc	tccgttgcate	tcaccttgc	cctgcggaga	720
aagtatccat	catgggtat	gcaatgggg	ggctgcatac	gtttgtatcc	gtctactgtc	780
cattcgacaa	ccaaaggcggaa	catcgatc	agogacacg	tactcgatc	gaagccggtc	840
tttgtcatac	ggatgtatcg	gacgaaagac	atcagggggt	ccggccaggc	gaacttgtcg	900
ccaggctcaa	ggcgcgtatc	ccggacggcg	aggatctgt	cgtgaccat	ggcgatgtcc	960
gcttgcgc	tatcatgtg	gaaaatggcc	gtttttctgg	attcatcgac	tgtggccggc	1020
tgggtgtgg	ggggccatcg	caggacatag	cggttgcgc	ccgttatgt	gctgaagacg	1080
tttgcggccaa	atgggtgtac	cgttcttcgc	tggtttacgg	tatccgcgt	cccgatgtcc	1140
agccatcgcc	cttctatgc	cttcttgacg	agttttctgt	agttttaaaac	gaccacaacg	1200
gttttcctat	atggggatcc	atccggcccc	tetttcccttc	ccccctcta	cgttactgtc	1260
cgaaggccgt	tggaaataagg	ccgggtgtcg	ttttgtatca	tgttatttc	caccatattt	1320
ccgttcttttgc	gcaatgttgag	ggccggggaaa	cttggccctg	tcttcgtac	ggggatcttc	1380
agggttgc	cccccttcgc	caaaaggatc	caagggttcgt	tgaatgtcg	gaaggaaagca	1440
gtttccttgc	aaagtttgc	aagacaaaca	acgtgttgc	cgacccttg	caggcagcgg	1500
aacccttcgttgc	ggggccacac	gttgcgttgc	ggccaaacgc	caatgtata	agatataatcc	1560
gcaaaaggccc	cacaacccccc	gttgcgttgc	gttgcgttgc	tagttgtta	aaagatcaaa	1620
ttgggttctct	caacgcgtt	caacaagggg	ctgaaaggatc	ccccggatgt	accggatctgt	1680
atgggatotc	atctggggcc	tcgggtcaca	tggtttatcc	gttgggttgc	gagggtttaaa	1740
aacgtctagg	ccccccggaa	cacggggggac	tggtttttct	tttggaaa	cgataatacc	1800
atggggcota	tttacggcata	cttcccaacag	acggcgaccc	tacttgcgt	catcatccat	1860
agccatccac	ggccggccac	ggggccatcg	ggggggggac	tccaaatgtt	cttcacccgca	1920
acacaatctt	tcttggccac	ctggcgttac	ggcggtgttt	ggatctgtca	tcatgtgtcc	1980
ggttcaaaatc	cccttgcgg	ccaaaggccc	caaatccatcc	aaatgtacac	caatgtggcc	2040
caggactctcg	tcgggtggcc	aggggcccccc	ggggccgggt	cotttgacacc	atgcacccgtc	2100
ggcagcttgc	accctttacttgc	tcgggtcagg	atcgccgtat	tcattccgg	ggccgggggg	2160
ggcgacacca	ggggggggatc	actctccccc	aggccgggt	cottatgtaa	gggtcttccg	2220
ggcggttccac	tgctctgc	ctcgccccac	gttgcgttgc	totttgcgg	tgccgtgtgc	2280
acccgggggg	ttggggggatc	tttgcgttgc	tttgcgttgc	agttatgtgg	aaccactatg	2340
cgggtccccc	tcttcgttgc	caactctgc	cctccggccgg	taccggccag	atccggatgt	2400
gccccatccat	acggccatccat	ttgttagggcc	aaagacatcg	agggtgcggc	tcgcgtatgc	2460
gccccaaagggt	ataagggtgc	tgtctgtgc	ccgtccgtcg	ccggccaccc	agggttccgg	2520
ggttatgtatgt	ctaaaggccat	ttgtatgtgc	cttaacatca	gaacccgggt	aaggaccatc	2580
accacgggttgc	ccccatccat	gtacttcacc	tttgcgttgc	tttttgcgg	cggttgcgtc	2640
tcttggggggcc	cctatgtatc	cataatatgt	gtatgtgc	actcaatgt	cttcgaccat	2700
atcttggccat	tcggccatgt	cctggccaa	ggggacacgg	tttgcgttgc	actctgtgt	2760
ctggccacatcc	ctatgtcc	ggatgttgc	atcggtccgc	atccaaatcc	cgagggtgt	2820
gctttgtccat	gcactgttgc	aaatccccc	tatggaaaa	ccatccccc	cgagaccatc	2880
aagggggggg	ggggccatccat	tttttgcgtat	tcggaaaga	aatgtgtatg	gtctccggcc	2940
aaatgtgtcc	gcctcggtat	caatgtgtgc	gtatattacc	ggggccgttg	tgtatctgc	3000
ataccaacta	ggggagacgt	cattgtgt	gtcaacccgc	cttcaatgt	gggttttacc	3060
ggcgatccat	actaaatgt	gtactgtca	atcggtgtca	ccacacatcg	cgactgtcc	3120
ctggccacatcc	ccttccatcc	ttggggccat	ggggccatcg	acccgtccgc	gtcaccgtcc	3180
caggccggat	ggcaggatcc	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	3240
gaacccggccat	ggggccatccat	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	3300
gctttgttgc	agttttgtcc	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	3360
ccagggttgc	ccgtctgc	ggggccatccat	tttttgcgttgc	atccgggttgc	gactccggag	3420
accacatcat	acggccatccat	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	3480
ctgttgcgtat	accaggatccat	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	3540
caaatgtgg	atgttgcgtat	acggcttac	tttttgcgttgc	atccgggttgc	gactccggag	3600
tatagggttgc	ggagccgttca	aaacggagggt	actaccacac	acccctatata	caaataatcc	3660
atggcataatc	ttgtcgttgc	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	3720
gtcttagatgc	cttcggccatcc	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	3780
atcatcttgc	ccggaaaatcc	ggccatccat	cccgacgggg	aaatgtttta	ccggggatgtt	3840
gatgatgttgc	aaatgtgtcc	cttccacatcc	ccttacatcc	aacccgggtat	gtcggctcc	3900
gaacatccat	aaacggaggatc	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	3960
gtctgtccat	ccgttggatc	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	4020
tggaaatccat	tcacggggatc	aaatattatca	tttttgcgttgc	atccgggttgc	gactccggag	4080
gogtagatcat	cactgtgtcc	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	4140
acccttcgttgc	tttacatccat	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	4200
gctttgttgc	tcgttgcgtat	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	4260
aagggttgc	ttggatatttt	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	4320
tttaaggatccat	ttggatatttt	tttttgcgttgc	tttttgcgttgc	atccgggttgc	gactccggag	4380

atcccttc	ctggggccct	agtctgtcg	gtcgtgtcg	cagcgatact	ggctcgccac	4440
gtggggccag	gggggggggg	tgtgtcggtt	atggaaaggc	tatagcggtt	cgcttcgccc	4500
ggtaaccacy	tctccccccac	gactatcg	ctcgagacgg	tgatagcggt	4560	
cagatctct	ctagtottac	catcaatcg	ctgtgcggaa	acgtcgcaat	4620	
gaggactgt	ccacgocatg	ctccggctcg	tggcttaagg	gtttttggca	4680	
acgttgta	tctatttca	gacctggcgt	cagttcaaga	tcttcggccg	4740	
gtccccctt	tcttcgtca	acgtgggtac	aaggggatct	ggggggggca	4800	
caaaaaaccc	gccccatgtt	acccagacat	acccggatct	ccgcattatc	4860	
atcggtggc	cttaggacct	ttcccccggc	ccaaattttt	taacgctgtac	4920	
accacggggc	cttcggccgg	gttgcggcgg	gttggggggat	gtgggggtgt	4980	
gtgtcgagg	aatgtcggtt	cccgctgtac	gttgcggccg	cttacactcg	5040	
accactgaca	acgttaaagt	gtacgtccgt	gtgtgcggac	gacggggatcg	5100	
gatgggggtc	gttggccgtt	tcacatctgt	gttgcggaaa	ccttcctacg	5160	
acatccatgg	tccgggtca	tcacatctgt	gttgcggatc	cgagggccaa	5220	
ccggacgtag	cagtgtccat	tccatgtcgt	acccggatct	cccacatccat	5280	
gttaaegcta	gggttgcggat	ggggatccct	cccttcctgg	taacgctgtac	5340	
ctgtctgcgc	cttccttggaa	ggcaacatgc	actaccggct	atgactccccc	5400	
ctcatcgagg	ccaaactctct	gttggggccgg	gagatggggc	ggacatccat	5460	
tcagaaaaat	aggtagtat	tttgcgttct	tttgcggatc	cccgctgtgg	5520	
agggaatgt	cgttccgggg	ggagatctgg	ggggatccgg	gggatattcc	5580	
ccatataatgg	cacggccggg	tttacacccat	ccacttgcgt	tgccggatgt	5640	
taatgccttc	cagtgttaca	cggtgttcca	tttgcgttgg	ggacggccgg	5700	
cctccatgtt	ggaaaggaggc	gggttgcgtt	tttgcgttgg	ttccatgtcca	5760	
gatgtcgccg	caaaagcttt	ccggcgttcc	tttgcgttgg	atgactcgac	5820	
acggcccttc	cttggccggcc	ggccgttcc	tttgcgttgg	ccggcggccgg	5880	
tcccttcatgc	cccccttgcg	ggggatccgg	tttgcgttgg	gttgcgttgg	5940	
tctaccgtaa	gggggggggg	tttgcgttgg	tttgcgttgg	tttgcgttgg	6000	
acaggccgtt	tgatcacggc	tttgcgttgg	tttgcgttgg	tttgcgttgg	6060	
acaaacttcc	tgttcggctt	ccacaaatgg	tttgcgttgg	tttgcgttgg	6120	
ctggggcaga	agaagggtcac	ttttgcgttgg	tttgcgttgg	tttgcgttgg	6180	
gtgtcgaaagg	agatgtgggg	tttgcgttgg	tttgcgttgg	tttgcgttgg	6240	
gaacgcgttg	agtgcgtccg	ccacacatgg	tttgcgttgg	tttgcgttgg	6300	
gacgtccgg	acctatccag	tttgcgttgg	tttgcgttgg	tttgcgttgg	6360	
ctggaaagaca	ctggggccac	tttgcgttgg	tttgcgttgg	tttgcgttgg	6420	
gttcaaccag	aaaggggggg	tttgcgttgg	tttgcgttgg	tttgcgttgg	6480	
gttcgtgtt	ggggggaaaat	tttgcgttgg	tttgcgttgg	tttgcgttgg	6540	
atgggtttt	cataccggatt	tttgcgttgg	tttgcgttgg	tttgcgttgg	6600	
gtcttggaaag	cgaagaaatg	tttgcgttgg	tttgcgttgg	tttgcgttgg	6660	
acggctgtt	aaatgtacat	tttgcgttgg	tttgcgttgg	tttgcgttgg	6720	
ccggccggcc	aattggacat	tttgcgttgg	tttgcgttgg	tttgcgttgg	6780	
actaatttca	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	6840	
accaggctgt	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	6900	
aaatcccgagg	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	6960	
ggggggggcc	aaaggggggc	tttgcgttgg	tttgcgttgg	tttgcgttgg	7020	
tctggcccc	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7080	
tcctccaaatg	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7140	
gaccggccac	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7200	
tcctgggttag	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7260	
actcttattct	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7320	
acttaacgggg	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7380	
catggcccta	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7440	
tcatgtccca	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7500	
gtcccggtata	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7560	
aactggggcag	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7620	
ttatccatgtt	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7680	
ttatcaacgggg	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7740	
catggcccta	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7800	
ttttttttttt	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7860	
ttttttttttt	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7920	
tagtcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	7980	
agatcaagta	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	8040	
gcatatttgc	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	8100	
atgtttttatcc	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	8160	
ctcatcgta	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	8220	
gtactgtcgcc	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	8280	
otgtgttcgtat	tttgcgttgg	tttgcgttgg	tttgcgttgg	tttgcgttgg	8340	

tccgacccgt	ttggccgcgc	cccagtctgt	ctagttcgc	tacttggagc	cactatcgac	8400
tacggatca	tggcggaccac	accggctctg	tggatctct	acgcggagc	cattgtggcc	8460
ggcatccgg	ggcccaacgg	tgccgttgt	ggccgtata	tcccgocat	ccccatggg	8520
gaagatcg	ctggccactt	cgggctatg	acgcgttgt	tccgggtggg	tatggtggca	8580
ggcccgctgg	ccggggact	gttggccgc	atctcttcg	atcgaccat	ccttgcggcg	8640
gogggtctca	aacggctcaa	cttactact	ggttgttcc	taatgcagg	gtcgccataag	8700
ggagagcgt	gaccgtatgc	cttggagagc	ttcaaccag	tcagtcctt	ccgggtgggg	8760
cggggccgg	ctatcgctcg	cgcaacttgc	actgtcttt	ttatcatgca	actcgtagga	8820
cagggtccgg	cacgcgtctg	ggtcatttc	ggcgaggac	gctttcgctg	gaggcgacg	8880
atgatcgcc	tgtcggttc	ggtatttoga	atcttgcacg	ccctgcgtca	agccttcgtc	8940
actgtgtcc	ccacccaacg	tttcggggag	aaggaggcca	tatgcggccg	atggcgccgc	9000
gacggcgctgg	gtacgttctt	gttgggttc	ggccggcg	gctgtatggc	cttccccat	9050
atgatttctt	tcgtctccgg	cggtatccgg	atggcccg	tgcaggccgt	gtctgtccagg	9120
caggtagatq	aacggatccat	ggggacagct	caaggatcgc	tcggcgctt	taccagctt	9180
acttcgtatc	ctggacccgt	gatcgctcag	gcatgtttat	ccgcctcgcc	gagcacatgg	9240
aacgggttgg	catggatgt	aggccggcc	ctataccat	tctgcctccc	ccgcgttgtgt	9300
cgcgtgtcat	ggacggccgg	cacccgtcc	tgatggaa	ccggggccac	ctcgatcaacg	9360
gattcaccc	tccaaagatt	ggggccatc	aattttcg	gagaactgtg	aatgdgcaaa	9420
ccaaaccttgc	cgagaaacata	tccatcgct	ccgcacatc	cagcagccgc	acggggcgca	9480
tctccggcag	cttgggttcc	ttggccacgg	tgccgtatg	ctgtgtctt	tcgttggaga	9540
cccggttgg	cttgggggtt	ttggccatgg	tttgcattact	gttacgaga	tgaaatcccg	9600
gaacatgttgc	cgactgtc	tgcaaaaacgt	cttgcacccgt	aaacacaaca	tgaatggctt	9660
toegggttccg	tgtttcgtaa	agtcgtggaa	cgccggaa	agcgccttc	accattatgt	9720
tcccgatctg	catcgccgg	tgctgttgc	taccctgtt	aaacactata	tctgtttaa	9780
cgaaggcgctg	gtatgttcc	tgatgttttt	tttctgttgc	ccggccgtatc	cataccggcca	9840
gttggtttacc	ctcacaaacgt	tcggatcaac	ggccatgttcc	atcatcgta	accggatgtcg	9900
tgagcatctt	cttgcgttgc	atccgttata	tttacccat	aaacaaat	tcccccttat	9960
acggaggcat	caagtgtacca	aacaggaaaa	aaccgcctt	aaacatggcc	gctttatcag	10020
aaggccagat	ttaaacgtt	ttggggaaact	caacggatgc	gacggggatg	aaacggccaga	10080
catotgttgc	tgcgttccac	accacccgtca	tgatgttttac	ccgcgtcgc	tcggcgctt	10140
cgtgtatgc	ggtggaaaac	tctgacatc	tgatgttcc	gagacgggtt	cagetttgtt	10200
gtaaagggttgc	ggccggggatc	cgacggcc	tcaggccgc	tcaggccgtt	ttggccgggt	10260
tcggggccgc	ggccatgtaccc	agtcgtacgt	cgatgtccgt	gttgcggat	gtttactatgt	10320
ggccgtatcg	aggcgtatcg	actgtgttgt	cattatgtc	tgatgttgc	acccgcacaga	10380
tcgcgttaagg	gaaaataatcg	catcgccgc	tttccgttgc	cctcgatc	tgactcgctg	10440
cgctcggtcg	ttccgggtcg	gcccgggtt	tgatgttca	caaaggccgt	aatacggtt	10500
tccacagat	cggggggatc	cgacggaa	aatatgttgc	ggaaaaggcc	gaaaaaggcc	10560
aggaaacccgt	aaaaggccgc	gttgcgttgt	ttttccatca	ggctccggcc	ccctcgacag	10620
catcacaat	atcgccatcg	aaatgttgc	tgccggaa	cgacaggatc	ataaaagatac	10680
cagggttgc	ccctgttgc	ttccgttgc	cgcttcctc	ttccgcaccc	ggccgttacc	10740
ggataccgtt	ccgcctttt	cccttcggga	agcgtggcc	tttccatag	ctcacgttgt	10800
aggtatccat	tttccgggtt	tttgcgttgt	ttcaacgttgc	gctgtgttgc	cgaaacccccc	10860
gttcaggcccg	accgcgtcg	cttacccgtt	aactatgtc	tttgcgttca	cccggttaa	10920
cacgaacttat	ccgcacttgc	agcaggccat	tttgcgttca	tttgcgttgt	gagggtatgt	10980
ggccgtgtctt	caaggttctt	tttgcgttgt	tttgcgttgt	tttgcgttgt	tttgcgttgt	11040
tttgcgttgt	tttgcgttgt	tttgcgttgt	tttgcgttgt	tttgcgttgt	tttgcgttgt	11100
tccggccaaac	aaaccacccgc	tttgcgttgt	tttgcgttgt	tttgcgttgt	tttgcgttgt	11160
cgcgaaat	aaaggatcttca	tttgcgttgt	tttgcgttgt	tttgcgttgt	tttgcgttgt	11220
tggAACAAA	acttcacgtt	tttgcgttgt	tttgcgttgt	tttgcgttgt	tttgcgttgt	11280
ttatccaaat	tttgcgttgt	tttgcgttgt	tttgcgttgt	tttgcgttgt	tttgcgttgt	11313

<210> 11
<211> 11184
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 11

ggccgcggcc	gattggggcc	gacactccac	catagatca	tccctgttgc	ggaactactg	60
tcttcacggca	ggaaaacgtct	agccatggcc	ttatgttgc	tgttgttgc	cctccaggac	120
ccccctcc	ggggagacgc	ttatgttgc	ttatgttgc	ttatgttgc	ttatgttgc	180
gacggccgg	ttatgttgc	ttatgttgc	ttatgttgc	ttatgttgc	ttatgttgc	240

gcttcgttgc	tgcgtggcgc	cggcatcgct	ggagcggtct	ttggcagcat	aggccctggq	4260
aagggtgttgc	tggatatttt	ggcagggtat	ggagcagggg	tggcaggegc	gtctgtggcc	4320
ttaaaggcc	tgagggggaa	gatccccc	acaggagacc	tggttaactt	actccctgt	4380
atccctcc	ctggccccc	agtgcgttgg	tgctgtgtgg	caggcatat	ggtgtggcac	4440
gtggggcc	gggggggggg	tgtgcgttgc	atgaacccgc	tgataggtt	cgcttcgggg	4500
ggtaacca	ttcccccc	gactatgttgc	cgtggagagc	acgtcgacgc	acgtgtca	4560
catactcc	ctagtttt	catcaactcg	ctgtgtga	ggcttaccca	gtggatcaac	4620
gaggactgt	ccacggccat	ctccggctcg	tggctaagag	atgtttggga	ttggatgtc	4680
acgggttgttgc	ctgttttt	gactcgctc	cagtccaa	tcctcgccgg	atggccggga	4740
gtcccttctt	tclcatgtca	acgtgggtac	aaggaggtt	ggccggggga	eggcatatcg	4800
caaaccac	ggccatgttgc	agacagatc	acccggat	tgaaaaacgg	ttccatgagg	4860
atcggtgggg	ctggacgttgc	tagtaacatgc	tggcatggaa	cattccccc	taacggctac	4920
accacggcc	cttcggccgc	cttccgggg	ccaaaatttt	tcaggggct	gtgggggttg	4980
gtgtgtgg	agtaatgttgc	gggttacggg	gttggggat	tccactatgt	gacggcgatc	5040
accactgaca	acgttaatgt	cccggttgc	gttccggccc	ccgaattttt	cacagaatgt	5100
gatgggggtgc	gggttgcac	gtacgttcc	gggtgttca	coctcttca	ggaggaggatc	5160
acatccctgg	tcgggttcc	tcataactctg	gttgggttca	agctccat	egagccggaa	5220
ccggacgt	caatgttgc	ttccatgttgc	acccggat	ccacatttac	ggccggggac	5280
gctaaggta	gggttgcgc	ggatcttcc	cccttcttgc	coacttac	atgtatccag	5340
ctgtgtgc	cttccatgttgc	ggcaacatgt	actaccggc	atgttccccc	ggacgttgc	5400
ctatcgaga	cccaatcc	gttggccgc	gatgggggg	ggaaatccat	ccggatgtgg	5460
tcagaaaaata	aggtaatgt	ttttgtatgt	tttgcgttgc	tcggacggc	ggaggatgtac	5520
agggaaat	ccgttccgc	ggagatcttgc	cgagggttca	ggaaaatttcc	tcgagcgat	5580
ccatcatggg	cactccgg	tttacacatgc	ccactgttag	atgttgcggaa	ggaccggcc	5640
taatgttcc	cgtgttgc	cggtgttca	tttgcggcttgc	ccaaaggccc	tcggatccca	5700
cctccacgg	ggaaaggac	gttggtctgt	tcagatgttgc	tcgtgttcc	tgcccttggg	5760
gagatcgcc	aaaaggatc	ccggatgttgc	gaatctgttgc	tcgtgttgc	ccggacacgg	5820
acggcccttc	ctgttgc	cgtgtgttgc	tgctgtatgt	otacatcat	gacaggcgcc	5880
ctgtatcgcc	catcggtgc	ggggggaa	aaatgttcc	tcataatgt	gagaacttct	5940
ttgttccgc	accacaaat	gggttctatgt	atcaatccat	ggggccgg	ctgtggggat	6000
aagaaggta	cccttgcac	actcgatgt	ctggatggcc	actacccgg	cgtgttca	6060
gagatgttgc	ccaggccgtc	caagatgttgc	gttacatgttgc	tatccgttgc	ggaaggctgt	6120
aagttgttgc	cccccacat	ggccatgttgc	aaatgttgc	atggggcaaa	ggaeatccgg	6180
aacctatcca	ggcaaggccgt	taaccatc	gttggccgttgc	ggggggactt	gttggaaatc	6240
actgttgcac	caattgttgc	cattatgttgc	gttttttttt	cgatccaa	ggggggatgt	6300
gagaagggg	ggcgcaac	agtcgcgttgc	atgttatcc	catgggttgg	gggttgcgtg	6360
tgcgtggaaa	ttggcccttta	cgatgtgttgc	tcacccatc	tcaggccgt	gttggcttct	6420
tcataatcg	tcataatcg	tccgttgcac	cgggatgttgc	tcgtgttgc	tgctgtggaa	6480
gccaatgttgc	ggccatgttgc	tcggccat	gacaccgttgc	tttttttttt	tttttttttt	6540
gagaatgttgc	ttccgttgc	ggatgttgc	taccaatgttgc	tttttttttt	tttttttttt	6600
agacaggcaca	taagggttgc	cacatgttgc	ggggcccttgc	gactaatttt	ggggcccttgc	6660
aaagggggg	actcggttgc	tcggcggttgc	cgccgggggg	gttgcgttgc	gaccagctgc	6720
ggtaataatcc	tcataatgttgc	tttggggat	tcggcggttgc	gttgcgttgc	gaagttccat	6780
gactgttgc	tgtgttgc	cgggatgttgc	tttgggttgc	tttgggttgc	cgccgggggg	6840
caagaggatc	ggggggat	acggggatgttgc	tcggatgttgc	tttttttttt	tttttttttt	6900
cctggggatc	cgcccaatcc	aaatgttgc	tttgggttgc	tttttttttt	tttttttttt	6960
gtgtcaatgc	cgacatgttgc	atgttgc	tttttttttt	tttttttttt	tttttttttt	7020
accccttc	cgccgggttgc	tttggggat	tcggatgttgc	tttgggttgc	tttgggttgc	7080
ggcaatccat	tcataatgttgc	ggccacatc	tttggggat	tttgggttgc	tttgggttgc	7140
tttccatcc	tttccatcc	ggcaacatc	tttggggat	tttgggttgc	tttgggttgc	7200
ggctgttgc	ttatgttgc	actgttgc	tttggggat	tttgggttgc	tttgggttgc	7260
agegtatcc	caatccat	ttatgttgc	tttggggat	tttgggttgc	tttgggttgc	7320
aggaaatcc	gggttaccc	tttggggat	tttggggat	tttgggttgc	tttgggttgc	7380
aggctactgt	cccgagggg	tttggggat	tttggggat	tttgggttgc	tttgggttgc	7440
gttaaggcc	aggccaaat	cactccatc	tttggggat	tttgggttgc	tttgggttgc	7500
tgggttgc	ctgtgttgc	tttggggat	tttggggat	tttgggttgc	tttgggttgc	7560
cgctgttgc	tgtgttgc	tttggggat	tttggggat	tttgggttgc	tttgggttgc	7620
aaccgttgc	ccccggat	aaatccatc	tttggggat	tttgggttgc	tttgggttgc	7680
ttttttttt	ttttttttt	ttttttttt	ttttttttt	tttgggttgc	tttgggttgc	7740
ttttttttt	ttttttttt	ttttttttt	ttttttttt	tttgggttgc	tttgggttgc	7800
aaggctccgt	aggccgttgc	tttggggat	tttggggat	tttgggttgc	tttgggttgc	7860
actctgttgc	ggccggccat	tttggggat	tttggggat	tttgggttgc	tttgggttgc	7920
cgacccaaat	tttccatcc	tttggggat	tttggggat	tttgggttgc	tttgggttgc	7980
cgtttaatcc	tttccatcc	tttggggat	tttggggat	tttgggttgc	tttgggttgc	8040
atctgttgc	ccgttccatcc	tttggggat	tttggggat	tttgggttgc	tttgggttgc	8100
ggcccttgc	gggtatgttgc	tttggggat	tttggggat	tttgggttgc	tttgggttgc	8160

ggcgtatatg cgttcatatgc atttctatgc gcaccggc ttccgggact ctggcggaccc 8220
 ttggccccc gccccagtct gtcgcttcg ctacttgag ccactatcg ctaegcgatc 8280
 atggggaccc caccggctc ttggatctc tacggcggac goatcggtgg cggccatcc 8340
 ggcccaacag gtgggttgc tggggctat atccggcaca tcaccggatg ggaagatcg 8400
 gtcgcactat tcgggctat gagcgctgtt ttccgggtt gatctggcgc aggccccgtg 8460
 gccggggaccc ttggggcgc catctttcg catgcacat tccttgccg ggcgggtgtc 8520
 aacggccctt acctactat gggtcttgc ctaatcgagg agtcgcataa gggagagcgt 8580
 cgaccgtgc ccttgagagc ctccaaacca gtcagcttc tccgggtggc gccggggcatg 8640
 actatcgcc cccgacttt gactgtttc ttatcatge aactcgtagg acagggtggc 8700
 gcagcgctt gggctattt egggaggac cgcttcgtt ggagcgcgac gatgtatggc 8760
 ctgtcgcc cgggtattcg aatcttcgc acgttcgtc acgttctgt cactgtgtcc 8820
 gccaacaaac ttggccggc gaaggcggc attatcgcc gcatggggcc cgacggcggt 8880
 ggtactatc ttgtgggtt cggcgatgg gcttcccat tatgtatctt 8940
 ctggcttcg gggccatgg tgccggccg ttccggggc tgctgtccg gcaaggatgt 9000
 gacgaccatc agggacactt tcaggatcg ctccgggtc ttaccaggct aacttcgtc 9060
 actggacccg tgatcgatc ggggatattt gcccgttcgg egagcacatg gaacgggttg 9120
 gcatggatc tagggccggc ctatacttc tggtctcctc eccgtgtgg tcgggggtca 9180
 tggagccggg ccaactcgac ctgaatggaa gccggggca ctccgtcaac ggatccacca 9240
 ctccaaagaat ttggaccaat caatttttcg ggaaactgt gaatcgcaaa accaaacctt 9300
 ggcagacatat atccatcgcc ttccggatcc ctccggccg caccggggcc atctcgggca 9360
 gcttggggcc ctggccaccc tgccgtatg tgctgttccct gtctgttggg accccggatcg 9420
 gctggggggg ttgtctattt gtttagcata atgaatccatc gatacgccg cgaacgttga 9480
 gegaactgtc ctggaaaacg tctgcgaccc gggcaaaac atgaatggtc ttccgggttc 9540
 gtgttctgtt aagtctggatc acggcgaaatg cggccgttc caccattat ttccggatct 9600
 gcatcgccat atgtgttgtt ctacccttgc gaaacactat atctgttata acgaaggcgt 9660
 ggcattgtacc ctgatgtttt ttctctgtt ccggccgtt ccgttccggc agttgtttac 9720
 cctcaacaacg ttccatgttac cggggatgtt catatcgatc aaccggatcc tgtagatcc 9780
 tctctgtttt catcggtatc attacccttca tgaacagaaaa ttccccccta caccggaggca 9840
 tcaatgtacc aaacagaaaaaaacccctt taaatcgcc cgttattatca gaagcccgac 9900
 attaaacgttcc tggggaaacatc tcaacagatg ggacggcgat gaacaggcag acatctgtca 9960
 atcgcttccat gaccatcgatc atgatgtttt cccggatgtc ctccgggtt tgctgtatga 10020
 cggtggaaatcc ctgtgacaca tccgtatcc gggatcgatc acgttgcg tggatggc 10080
 tggggggggc agacaaaggcc gtcaggggccg tgccgggggtt gtcggggggc 10140
 agccatgacc cgtatcgatc gggatcgcc agttgttccat ggtttacta tgccggatcca 10200
 gagccatgtt tactgttgc gacccatgtt cccgtgttgc taccggccatc atcgatgttgc 10260
 agaaaaatacc gcatcaggccg ctccctcgat tccctcgatca ctgactcgat ggcgtcggtc 10320
 gttccgggtcc gggggggatc atcgatgttcc tccaaaggccg taatatcggtt atccacagaa 10380
 tcaggggatcc acggcggaaaat gaaatcgatca gggaaaggccg acggaaaggcc cggggatgttgc 10440
 aaaaaggccg cttgttccat gttttccatc aggttccggc cccctgtacca gcatcacaat 10500
 aatcgacccat ccaatgttgc tggttgcggatc ttccggatcc tataatggatata ccaggcggtt 10560
 cccctgttca gtccttcgtt gggcttcgtt gttccggatcc tgccgttac cggatccctg 10620
 tcccgcttcc tcccttcgtt gggatcgccg ctccatcgatc gtttccatc gtcacgtgt taggttatcc 10680
 agtctgggtt aggttccgtt ctccaaatgtt ggttgcgttgc acggacccccc cgttccggatcc 10740
 gacccgttgc ctttccatcgatc taatatcgatc ttccggatcc acccggatgg acacggactta 10800
 tggccacttgc cggccatccat tggttacaggcgtt attagcaggc acggatgttgc agggcggtt 10860
 acaggatgtt tggatgttgc gtcctactat ggttactatc gggatggatc atttggatcc 10920
 tgccgtatcc tggggatgttccatccatc gggatggatc gtttccatcgatc gtttccatcgatc 10980
 caaacccatcc ctggatgttgc tggttccatcgatc gtttccatcgatc gggatggatc 11040
 aaaggatctc aagaatgtcc ttgtatccatc tccacggggatc ctggatcgatc gtttccatcgatc 11100
 aacttcgtt aagggtttt ggttcatgtt gttatccatcgatc ggttccatcgatc ctggatccatc 11160
 ttctatgttac tccgtatcgatc tata 11184

<210> 12
 <211> 11313
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid

<400> 12

gccggccccc gatggggggc gacactccac catagatcac tccctgtga ggaactactg	60
tcttcacggc gaaaggctct agccatggcc tttagatgt tgctgtgtc cctccaggac	120
cccccttccc gggagggcca tagtggctcg cggaaagggt ggttacaccg gaattggccag	180

gacgaccggg	tccttttgc	gatcaacccg	ctcaatgcct	ggagatttg	gcgtcccc	240
gcgagactgc	tagccgaga	gtgttgggtc	gcaaaggcc	ttgtgtact	gcctgtatgg	300
gtgttgcga	gtggccggg	aggcttgcga	ggccgtgcac	catggcag	aatccataac	360
ctcaaaaga	aaccaaaa	cgcggccat	ttaacaaca	tggatgcac	gcagggttc	420
cggccgttgc	ggttgagagg	sttacgcgt	actgaatggc	aaacacaga	atcggtctgt	480
ctgatgcgc	cgtgttccg	ctgtcagcgc	aggggccccc	ggttttttt	gtcaagaccc	540
acccgtccgg	tcggccgt	gaaactgcagg	acggaggcag	gccccatcg	tgggtggcca	600
cgacgggggt	tccttgcgc	gtgttgcgtc	acgttgtcac	tgaaggccgg	aggggactggc	660
tgtatgggg	cgaatggcc	ggccaggatc	tctctgtatc	tcacccgtgt	ctggccggaga	720
aagtatccat	catgggtat	gcaatgcggc	ggctgcatac	gtttgtatcc	getacccgtcc	780
cattgcacca	ccaaaggc	catgcgtatc	acggaggcag	tactggatgt	gaagccggcc	840
tttgtatca	ggatgtatc	gcaagaaaggc	atcagggtgt	cgccocacgc	gaactgttcc	900
ccaggctca	ggccgcgtat	ccccggccgg	aggatctgt	ctgtaccatc	ggcgatgtct	960
gtttccgg	ttatcatgg	gaaatgggg	gttttttcgt	ttatcatgcac	tgtggccggcc	1020
tgggtgtggc	ggaccgtat	caggacatag	cgttgcatac	ccgtatgtt	gctgaagaggc	1080
ttggccgggg	atgggggtac	cggttccgtc	tgcttttacgg	tatccgggt	ccggatccgtc	1140
ageccatgc	ctttatccatc	ttttttgtatc	aggtttttcgt	atgttttttt	gaccacaaacg	1200
gtttccctct	agccggatca	atccggccccc	tttcccttcc	ccccccctaa	cggtactggc	1260
cgaaggccgt	tggataaagg	ccccggccgg	tttgcgtatc	tgttattttt	caccatattg	1320
ccgttccgtt	gcaatgttag	ggccggggaa	cttggccctc	ttttttgtac	gaggattctct	1380
agggttctt	cccttcgtcc	caaaaggatc	caagggtctgt	tgaatgtctgt	gaaggaaacgg	1440
gttctcttgc	aaagggttcc	aaacaaaca	acgtttgtatc	cgaccctttt	caggcaggccg	1500
aaccccccac	ctggcgacag	gtgcctctgc	ggccaaaaacg	cacgtgtata	agatacacct	1560
gcaaaaggcc	cacaacccca	gtggccatgt	gtgagttgg	tatgttgg	aaaggttac	1620
tggdtcttc	caaggatgtt	caaaagggg	cttggatgtt	ccccaaagggt	acccccatgt	1680
atgggtatcg	atctggggcc	tcggcgtaca	tgctttatcc	gtgtttatgc	gagggtttaaa	1740
aacgtttag	ccccccggaa	caagggggg	tttttttct	tggaaaaaca	cgatataacc	1800
atggcccta	ttacggcc	ctcccaacag	acggaggccc	tactttgt	catcatcaact	1860
agccctacac	ggccggggacag	ggaaaggcgt	gggggggggg	tccaaatgg	ctccaccggca	1920
acacaatctt	tcttgggg	ctggcgat	ggccgtgtt	ggactgttca	tatgttggcc	1980
ggctcaaga	cccttgcgg	ccccaaaggcc	ccaaatccac	aaatgttacac	aatgttgcac	2040
caggactcg	tcggctgtcc	aggccccc	ggggccgtt	catttttttt	atgcacccgtt	2100
ggcagctccg	accccttact	ggtcacgggg	catcgatgt	tatccgggt	ggccgggggg	2160
ggcgcacccac	ggggggggct	acttccccc	ggccgggtt	tatctttttt	gggttcttcg	2220
ggcggtccgg	tgttctccgc	cttggggcc	gtgttggca	tcttttccgg	tgccgtgtcc	2280
accccgagggg	tttgcgttgc	gggtggactt	gttccctgtc	agtttatgtt	aaccactatg	2340
cggcccccgg	tttccatccgg	caactcttc	cttccggcc	tttttttttt	atttccagggt	2400
gccccatctatc	acggccctatc	tttttttttt	tttttttttt	tttttttttt	tgctgtatgc	2460
gccccaaagggt	ataaagggtt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2520
gcttatgtt	ctaaaggccaa	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2580
accacgggtt	ccccccatcc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2640
tcttgggggg	ccatctatgt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2700
atccctgggg	tcggccatgt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2760
ctggccatcc	ctaaaggccatc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2820
gctttgttcc	gcaatgggaa	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2880
aagggggggg	ggccatccat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2940
aagggttgcg	ggccggact	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3000
ataccatca	ggccggacatc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3060
ggcgattttcg	actcaggat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3120
ctggacccgg	ccttgcaccc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3180
caggccgggg	ggccggact	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3240
gaacggccct	ccccggatgtt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3300
gctttgttgc	actgttccgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3360
ccagggttgc	ccgttgcgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3420
acccacatag	acggccatcc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3480
ctggtagat	accaggat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3540
caaatgttgc	agtgttccat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3600
tataggcttgc	gaggccgttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3660
atggatcatgc	tgttccatgt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3720
gttcttagat	cttctggccg	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3780
atccatgttgc	ccggaaaggcc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3840
gatgatgttgc	aaaggatgtgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3900
gaacatttca	aaacaaaggat	tttttttttt	tttttttttt	tttttttttt	tttttttttt	3960
gctgtccccc	ccggggatgtt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	4020
tgttccatgttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	4080
gctatgttgc	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	4140

accctcttgt	ttaacatct	ggggggatgg	gtggccggcc	aacttgc	tccaggcgct	4200
gcttttgtt	tcgttagggc	cgccatcg	ggagggcg	ttggcagcat	aggccctggg	4260
aagggtgtt	tggtatattt	ggcaggat	ggaggggg	ttggcagggc	gtctgtggcc	4320
ttaaaggct	tgacgggg	gacggcc	acggaggac	ttgtttaact	actccctgt	4380
atccctcc	tcggccgc	agtctgtt	ttgtctgg	cggcataact	ggctggggac	4440
gtggccccc	gggggggg	ttgtcagt	atgaacccgg	tgatagctt	cggttcggcc	4500
ggtaaacacg	tctcccccc	gcaactatgt	ctggagagcg	acgctgc	acgtgtca	4560
cagatccct	ctatgttt	catcaactcg	ctgtcgaaa	ggttcaacc	gtggatac	4620
gaggactgt	ccacggcc	ctccggctcg	ttggctaaag	atgttggg	ttggatatgc	4680
acgggttgtt	ctgttgc	gacggctgc	cgatccaa	ttccggccgg	atttgcggg	4740
gtccccctt	tctatgtt	acgtgggtt	aaggggat	ggcggggcc	eggcatat	4800
caaaccact	gccccatgtt	agcacagatc	acggacat	tgaaaaacgg	ttccatgagg	4860
atcggtgggtt	ctaggactt	tgtaaacacg	ttggcatgg	catttccat	taacggctac	4920
accacgggg	ctcgacgg	ctccgggg	ccaaattat	ttaggcggt	gtgggggtt	4980
gctgtgtt	ggatgtgtt	ggttacgggg	gtgggggatt	tccacta	gacgggcgt	5040
accactgaca	acgttaaagt	cccggttc	gttccggcc	ccgaattt	cacagaatgt	5100
gatgggtgtt	gggtgtc	gtatgc	ggtgtcaaa	cocttct	ggaggagg	5160
acattctgg	tcgggttca	ttcaatact	ttgtgggtac	agttccat	cgagccggaa	5220
ccggactgt	cgtgtctc	ttccatgtt	acggaccc	ccacattat	ggcgagac	5280
gctaaggta	gggtggcc	gggatctcc	cccttctt	ccatgttac	agctagcc	5340
ctgttgc	cttcgttgc	ggcaacatgt	actacccgt	atgtactcc	ggacgtgtac	5400
ctatcgagg	ccaaacctt	ttggcggc	gagatgggg	ggacatcac	ccgggtgg	5460
tcagaaaaat	aggtagtat	ttttggact	tttgcggc	ttcaagg	ggggatgt	5520
agggaaatgt	ccgttccgc	ggatgtct	ccgggttca	ggaaatttcc	tcgagcgt	5580
cccatatggg	cacggccgc	ttacaaccc	ccactgtt	agtcctggaa	ggaccggac	5640
taacgtcc	cagtgttac	ccgggttca	tttgcggcc	ccaaaggccc	ttcgatata	5700.
cctccaccc	ggaaagggg	gttgcgtt	ttggaaat	ccgtgttcc	tgccgttgg	5760
gagctogca	caaaggac	ccggacgt	gaatcg	ccgtgcac	gggcacgg	5820
acggcccttc	ctgaccc	cttccgac	ggcgcgg	gttgcacgt	tgatgtgt	5880
tcctccat	cccccttgc	gggggggg	ggggatcc	attcagaa	ccgggttgg	5940
tctaccatgc	gggggggg	ttatgtgg	gtcgtgtgt	gtcgtatgc	ctacatctt	6000
acaggccccc	tgatcagcc	atgcgttgc	ggggaaacca	agtcgtccat	caatgtact	6060
agcaactttt	tgtccgtc	ccacaaatcg	gtctatgt	caacatct	caggcaaa	6120
ctggggc	agaaggat	ctttgtac	ctgcagggt	ttggacacc	ctaccgggg	6180
gtgtcaagg	aggtgtgg	gggggttgc	acagttt	ttttactt	atccgtgg	6240
gaaggctgtt	atgtgtgg	ccccatct	ggccatcta	attttggta	ttggggcaaa	6300
gacgtccgg	acatccat	ggggccgtt	aaacatcc	gtccgtgt	gaaggactt	6360
ctggaaagaca	ctgagacac	aattgacacc	accatcat	aaaaaaat	gtttttctt	6420
gtccaaacc	agaagggggg	ccggcaacca	gtcgtcc	ttcgatattt	agatgttggg	6480
gttgcgtgtt	ggggaaaat	ggcccttac	gtatgtt	ccacccctt	tcaggcggt	6540
atgggttctt	catacggtt	ccaaatctt	cttggacac	gggttgcgtt	ctgtgtt	6600
gcttggaaag	cgaaagaat	ccctatggc	tttcgtat	acacccgt	ttttgtact	6660
acggtactt	agaatgtac	ccgtgttgc	gagttt	accaatgtt	tgacttgg	6720
cccgaaatca	gacaggccat	aaatgtgtc	acagacgg	tttacatgg	ggggccccct	6780
actaattttt	aaaggccgg	ctggcgctt	ccggcg	ggcgaggac	tgtaact	6840
accagtc	gtataatcc	ccatgttac	tttgggg	otggcgct	tcgagcttgc	6900
aagtc	actgcacat	gttcgtat	ggagacac	tttgggtt	ctgtgttgc	6960
ggcgccgg	aaaggac	ggggcgact	ccggcttca	ccgggtt	gactatgtac	7020
tctggggcc	ctgggggg	ggccaaacca	aaatgtac	gttggatgt	aatcatgtc	7080
tctccat	ttgtcgttgc	gtacgtatc	tttggcaaa	gggtgtact	tctccatgt	7140
gacccca	cccccttgc	gggggttgc	ttggagac	ctagacac	ttccgttca	7200
tccgttgc	ggcaatcat	catgttgc	ccaccc	ggcaaggat	gtatctgt	7260
actattttt	tctccat	tttgcgtt	tttttt	aaaaaaat	agatgttgc	7320
atctacgggg	cgtgttact	cattgtac	tttgcac	ttttttt	tttttttt	7380
catggcc	ggccat	ttttttt	ttttttt	ttttttt	tttttttt	7440
tcatgttca	ggaaactt	ttttttt	ttttttt	ttttttt	tttttttt	7500
gtccgtgtt	ggatgtat	ttttttt	ttttttt	ttttttt	tttttttt	7560
aactggcc	taaggccaa	ttttttt	ttttttt	ttttttt	tttttttt	7620
ttatccatgt	gggtgttgc	ttttttt	ttttttt	ttttttt	tttttttt	7680
ggccgaccc	gttgggttgc	ttttttt	ttttttt	ttttttt	tttttttt	7740
ctatccccc	accatgtac	ttttttt	ttttttt	ttttttt	tttttttt	7800
tttccctt	tttttttt	ttttttt	ttttttt	ttttttt	tttttttt	7860
ttttccct	tttttttt	ttttttt	ttttttt	ttttttt	tttttttt	7920
tagtgcgtt	agggtccgt	ttttttt	ttttttt	ttttttt	tttttttt	7980
agatcgttac	ctctgttgc	ttttttt	ttttttt	ttttttt	tttttttt	8040
gatatttgc	gaccat	ttttttt	ttttttt	ttttttt	tttttttt	8100

atgtttatcac	agttaaaattg	ctaacgcagt	caggcacccgt	gtatgaaatc	taacaatgcg	8160
ctcatcgca	tcctcgccac	cgtcacccgt	gatgtgtcg	gcatacggtt	ggtttatgcgg	8220
gtactgcgg	gcctcttgcg	ggatatacg	Cattccgaca	gcatacgccag	catactatgc	8280
gtgtgtcg	cgttatatgc	gttgatgcga	tttctatgcg	caccgttct	cgagacactg	8340
teccagccgt	ttggcccccgc	ccccatctcg	ttctgttcg	taacttggac	cactatgcac	8400
tacgcgatca	ttggcggaccc	accgcgtctc	ttgatcttc	acgcggacgc	cactgtggcc	8460
ggcatcaccg	ggccaccaagg	tgccgttgtt	ggccgtata	tcggccgatc	cacccgtatgg	8520
gaagatcggg	ctgcgttcgt	cgggttcatg	agcgttgcgt	tcggcggtgg	tatgttgcca	8580
ggccccgtgg	ccgggggact	gttggccgc	atctcttc	atgcacattt	ccttcggccg	8640
gggggtgtca	acggccatcaa	cattactactg	ggctgttcc	taatgcggga	gttcgcataag	8700
ggagagcgtc	gaccgtatgc	cttggagac	ttcaacccgc	tcagtcctt	ccgggtggcc	8760
cggggcatgc	ctatcgatgc	cgcactttag	actgtcttc	ttatcatgc	actcgtaga	8820
cagggtccgg	caagcgtctg	ggtcatttc	ggcggaggac	ggtttcgtg	gagccgcacg	8880
atgtatcgcc	tgtcgcttgc	gttattccga	atcttcgc	cccttcgttc	agccttcgtc	8940
actgttcccg	ccaccaaaacg	tttcggcgag	aaggaggcca	ttatgccegg	catggcgccc	9000
gacgcgttgc	gtacactgtt	gttggccgtt	gcgcacggag	gtcgatggc	cttcggccat	9060
atgtatccgtt	tcgtcttcgg	cggtatgggg	atgcggcgtt	tgccggccat	gctgttccagg	9120
caggtagatg	acgaccatcaa	ggacatttttc	caaggatcg	tcggggctt	taccaggctt	9180
acttcgtatca	ctggacccgt	gatcttcga	gogattttat	ccgccttcgg	gacacatgg	9240
aacgggttgg	catggatttg	aggccggccc	cttatacttc	tctgcctccc	cgcttgcgt	9300
cggcggtcat	ggacgggggg	cacctggacc	ptaatggaa	ccggggccat	ctcgcttaacg	9360
gattcacccat	tccaagaattt	ggggccatcc	aatttcgttg	gagaactgtg	aatgcgc当地	9420
ccaaacccttgc	ggagaatggat	ttccatcgct	cgccatcttc	cagcagccgc	acggggccca	9480
tctcggccag	cttgggggtcc	ttggccacccg	ttggccatgt	cgatgttcgt	tcgttggaga	9540
cccggtctgg	ctggcgccgg	ttgcctactt	tttagcagaa	tgaatcaccc	atacgccgac	9600
gaacgttgaa	cgactgtcg	ttccaaacgt	ttccgcaccc	aaacaaacat	tgaatgtgt	9660
tcgggtttcc	tgttctcgat	agttctggaa	ccggccggatc	ttccggccat	accattatgt	9720
tcccgatgttgc	catcgaggaa	tgctgttgc	taccctgttgc	aacacccata	tctgttataa	9780
cgaaggcgtc	ggatccggcc	ttgtgttttt	ttctctgttc	ccggccatcc	cataccggca	9840
gttgtttatcc	tcacaaacgt	ttccatgtttc	ggggatgttc	atcatcgat	accggatatc	9900
tgacgtatcc	cttcgttcc	atccgttata	tttcccccata	aaacagaaat	ttcccccattac	9960
acggaggccat	caagtgacca	aacaggaaaa	aaaccggccctt	aaatggccc	gttttatacg	10020
aagccagaca	ttaacgcctt	ttggagaact	caacggatgt	gacgcggatg	aacaggcaga	10080
catctgttgc	ttcgcttcac	accacgtgt	ttgtgttttt	ccgcgcgttcc	tcgcgcgtt	10140
cgggtatgttgc	gggtttttttt	ttttttttttt	ccgcgttcc	gagacgttgc	cagcttgcgt	10200
gtaaaggccat	ggccggggac	gaaacccggc	tttgcgggg	tcacgggttgc	ttggccgggttgc	10260
tcggggcgca	ggccatggcc	agtcacgtat	cgatagccgt	ttgttatactgc	gttttataact	10320
ggccgcattatc	agcagatgttgc	actggagatgt	cattatcgat	gggtgttataat	accgcacacg	10380
tgcgatggaa	ggaaaatatttgc	catcgccggc	ttttccgttgc	ccctcgatcc	tgactcgctg	10440
cgctcgatgttgc	ttccgttgc	ggccggggat	ttttttttttt	ttttttttttt	ttttttttttt	10500
tccacagaaat	caggggataa	cgccaggaaag	aacatgttgc	ttttttttttt	ttttttttttt	10560
aggaaaggta	aaaaaggccgc	tttgcgttgc	ttttttttttt	ttttttttttt	ttttttttttt	10620
catccacaaaaat	atcgacgttc	aaatcgatgg	ttttttttttt	ttttttttttt	ttttttttttt	10680
caggccatgttgc	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	10740
ggatacttgttgc	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	10800
aggatatctca	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	10860
gttcggatggca	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	10920
cacggactttat	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	10980
ggccgtgtatca	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	11040
ttttgttatcttgc	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	11100
tcccgcaaac	aaacccacccgc	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	11160
cgcagaaaaaaat	aaaggatctca	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	11220
tggaaacggat	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	11280
tagatcccttctc	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	ttttttttttt	11313

```

<210> 13
<211> 11184
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

<400> 13

```

gccggcccccc gattgggggc gacaatccac catagatcatc tccctgtgaa ggaactactg 60
 ttccatcgca gaaaggcgctt ageccatggcg ttagtgatcg ttcgtcgcc cttccggac 120
 ccccccctcc gggagacgca tagttttcttg gatcaacccg cggacccgtt gatgttccgg 180
 gacgacgggg tccttttttg agttttttttt cttcaatggctt ccgttgcggcc 240
 gogagactttt tagcccgatg agttttttttt gggatggggc ttgttgcgttact gctgtatgg 300
 tgcttttgcga gtggcccggg agttctcgta gacegtgcac catggacacg aatccataac 360
 ctcaaaagaaa aaaaaaaaggg cggccatggat tttttttttt tttttttttt tttttttttt 420
 cccggccgttg gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 480
 ctgtatggccg ogtgttccgg tttttttttt tttttttttt tttttttttt tttttttttt 540
 accttgtccgg tttttttttt gaaactcgagg tttttttttt tttttttttt tttttttttt 600
 ogacgggggtt tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 660
 tgctttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 720
 aatgtatccat cttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 780
 cttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 840
 ttgttgcatac gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 900
 cccaggcttac gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 960
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1020
 tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1080
 tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1140
 tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1200
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1260
 tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1320
 tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1380
 tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1440
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1500
 aacccccccaaat cttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 1560
 gcaaaaaagggg cttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 1620
 tggttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1680
 aatggatgttccg tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 1740
 aacgtcttagg tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 1800
 atggggcttccg tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 1860
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 1920
 aacccccccaaat cttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 1980
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2040
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2100
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2160
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2220
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2280
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2340
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2400
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2460
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2520
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2580
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2640
 tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2700
 atccatgttccg tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 2760
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2820
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2880
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 2940
 tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3000
 atccatgttccg tttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt 3060
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3120
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3180
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3240
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3300
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3360
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3420
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3480
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3540
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3600
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3660
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3720
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3780
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3840
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3900
 gttttttttt gttgttgcggc tttttttttt tttttttttt tttttttttt tttttttttt 3960

gtctgtcccg	cgggttggatc	caagggtggcg	acccttcgaa	ctttctgggc	gaagcataatg	4020
tggaaatttc	ctggccggat	acaatattta	ggaggctgtt	caacttcgtcc	tggccaaacc	4080
ggatcatgcat	cactgtggc	ataccacggc	tatccatcac	ggccgtccac	caccaccaat	4140
acccttcgtgt	ttaaatccat	ggggggatgg	ttggccgccc	aacttgcctc	tcccaagcgct	4200
gtttttgtgtt	tgttaggcgc	ccgcatecg	ggagccgtgt	ttggcagcat	aggcttgccg	4260
aaagggtgtt	ttggatattt	ggcaggatgt	ggagaaagggg	ttggcagcgc	gctgtgtggc	4320
ttaaaatggca	tgaggccggca	gtatccgtcc	accgggacc	ttgtttaacct	actccctgtgt	4380
atccctccccc	ctggggccct	atgtcgcc	gtcgtgtgg	cgacgatata	gcgtccggcc	4440
gtggggccag	gggagggggc	tgtcgastgg	atgaaacccgc	tgatagecg	cgcttcgggg	4500
ggtaaacccacy	tcccccacca	geactatgtb	ctgtgtggc	acgctgcgc	acgtgtcaact	4560
catagctctt	ctatgtttac	catcaatcg	ctgtgtgg	ggttccacce	gttgatcacac	4620
gaggactgtct	ccacccggat	ctccggctgc	ttggcataag	atgttgggg	tttgatgtcc	4680
acgggtgtta	ctgttattca	gaggcttgc	cgtccaa	tcttgcggcc	atggccggga	4740
gtcccccttc	tccatgtca	acgtgggtac	aaggaggatct	ggcgccggcg	ccgcacatcg	4800
caaaaacact	ggccatgtgg	ggacagatc	acggccatcg	tgaaaaacccg	ttccatgagg	4860
atcggtggc	cttggatcc	tagtaaactg	ttggatcg	cattccccc	atacgctc	4920
accacggggc	ctgtacgc	ctcccccgg	ccaaattatt	staggcgct	gtggcggtgt	4980
gtgtgtgtgg	agtaatgtgg	gtttagcgg	gtggggatgt	tcaactacgt	gcggccatcg	5040
accactgaca	tttttttttt	ccgggtgtgg	ttccggcc	cgaattttt	cacaagaatgt	5100
gatgggtgtt	gggttgcac	gtacgtccca	gggtgtcaac	cccttcatac	ggaggatgtc	5160
acatttttttt	tccgggttca	ttcaatctgt	ttgggttac	agcttccatc	cgagccggaa	5220
ccggacgttag	ctgggttca	ttccatgtac	accggccatc	ccacattac	ggcggagacg	5280
gctaatacggt	gggttgcac	ttccatgtcc	cccttcattgg	ccgtacgt	atgtatccatcg	5340
ttctgtgtgc	ttcttttttt	ggcaacatcg	ataccggcc	atgactcccc	ggacgtgtac	5400
ctcatcagggg	ccaaatccct	gtggggccgg	gagatggggc	ggaaatccatc	ccgcgtgggg	5460
tttagaaatata	aggtagtat	tttggatgt	tttggatcg	tttggatcg	ggaggatgt	5520
agggaaagtat	ccgttccgg	ggagatcc	ccggaggatcg	ggaaatttcc	ttcgagcgatg	5580
cccatatgtgg	cactttccgg	ttacaaatcc	ccacttgcgt	atgtctggaa	ggacccggac	5640
taatgtttttt	cagtgttaca	cggttgttca	tttgcgttcc	ccaaaggcc	ttccgttacca	5700
cctccacccgg	ggaaaggggc	gggttgcetg	tcgaatctca	cgtgttcc	ttccgttggc	5760
gaggtgttccca	caaaagccct	ccgcgttcc	taatgtcg	ccgtcgacac	ccgcacggcc	5820
acggcccttc	ctgtgttgc	cgatgttc	tttttttttt	catacatacg	gacaggccgg	5880
ctgtatcacgc	catgtgtgt	ggaggaaaaac	aaatgtccca	tcaatgtca	gaccaacttc	5940
ttgtccgtcc	accacactt	ggtctatgt	acaacatct	ccatcgac	ctcgccggag	6000
aaagggtgtca	cttttgac	atcgatggc	tttgcgttcc	actaccggg	cgtgtctca	6060
gagatgtttttt	cgaaagggtc	caacatgtt	gtatccatcg	ttatccgtgg	ggaaatgtgt	6120
aaatgtgtgtt	ccccccat	ggccgttcc	aaatttgtct	atggggcaaa	ggacgtccgg	6180
gaaatgttcc	gcaaggccgt	taaccatcc	ccgtccgtgt	atggggactt	gtgtggaaag	6240
actgtgtatcc	caattgttgc	cattatcc	ggaaaaatgt	atgtttttcc	ttgtccacca	6300
gtggggccatcc	ggccgttgc	atcgccatc	atcgatcc	cgatgtttcc	tttgcgttgc	6360
gggggggggg	tggcccttta	cgatgtgtt	tccacccctc	tcaggccgt	tttgcgttgc	6420
ttccatcgat	tccaaatctc	tttgcgttgc	ccgggtcgat	ttctgtgtt	tttgcgttggaa	6480
ggaaatgttttt	ggccatgtgg	tttgcgttgc	ttccatgtat	tttttgactc	ttccggggac	6540
gagaatgtaca	taatgttgc	ggatgttcc	tttgcgttgc	tttgcgttgc	ttccggggac	6600
aaaggccgtca	taatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	6660
ggatgtttttt	actgtgtgt	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	6720
ttccatcgatcc	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	6780
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	6840
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	6900
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	6960
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7020
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7080
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7140
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7200
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7260
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7320
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7380
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7440
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7500
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7560
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7620
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7680
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7740
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7800
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7860
ttccatcgat	ttccatcgat	tttgcgttgc	tttgcgttgc	tttgcgttgc	ttccggggac	7920

cgacccaaatt	ctcatgtttq	acagcttatac	atcgataa	ttaatgcgg	tagtttatca	7980
caggtaaatt	gctaaacgcg	tcaggcaccc	tgtatgaat	ctaacaatgc	gctcatgtc	8040
atccctcgaa	cgtcaccct	ggatgtgt	ggccatggct	tgttatgc	ggatgtcg	8100
ggcccttgc	ggatatgtc	ccatctccgc	agatccggca	gtcataatgg	ctgtgtgta	8160
gcccgtatag	cgttgatgtc	attttctatgc	gaccccggtt	tcggagact	gtccgcaccc	8220
tttggccgcc	gccccgttc	gtcgcttgtc	ctacttggag	ccatactcg	ctacgcgatc	8280
atggogacca	caccctgtct	gtggatctc	tacgcggac	gcatctgtgc	eggcgttacc	8340
ggcgcacacg	gtgggttgc	tgccgcctat	atcgccgaca	taaccggatgg	ggaaagatcg	8400
gctgcgttac	tcgggtctat	gagcgctgt	ttccgggtgg	gtatgtgtgg	aggccccctgt	8460
gcccggggac	tgttggggc	catctcttgc	catgcggccat	ttcttgcggc	ggcgtgtgtc	8520
aaacggctca	acactact	ggctgtcttc	ctaatgcagg	agtgcataa	gggagagatgt	8580
cgaccgtgc	ccttggagac	cttcaaccca	gtcagtcctt	tcgggtgggc	gccccggatgt	8640
actatcgcc	cggcactat	gactgttc	tttacatgc	aactctgttgc	acagggtgcgc	8700
gcagggcttc	gggtcatttt	cgccggagac	cgctttccgt	ggggccgcac	gatgtatcg	8760
ctgtcggtt	cggttattcg	aatcttgcac	gcccctgttc	aagcttgcgt	cactgtgtcc	8820
gcccaccaac	gtttccggca	gaaggcggcc	attatccggc	cgatggccgc	cgacccgcgt	8880
ggctactgt	tgtgtgggtt	cgccgcgcga	ggctgtgtat	cttccccatc	tatgtatttt	8940
ctcgcttccg	ggggcatcg	tgccgcggcg	tttgcggcc	tgctgttcac	gcagggtatgt	9000
gacgacatcc	aggcagacat	tcaaggatgc	ctccgggttc	ttacaggctt	aactctgtatc	9060
actggaccgc	tgatgttcac	ggcgatttt	gccccttcgg	cgagcacat	gaacgggttgc	9120
gcatggat	tgaggccgc	cttccatactt	gttcgttc	tcgcgtgtgc	tcgcgtgtca	9180
ttggacccgg	ccacctcgac	ctttagtggaa	gccccggggc	cttcgtctaa	ggatccacca	9240
ctccaagaat	tggagccat	caatttttc	ggagaactgt	gaatgcgc	accacccctt	9300
ggcagaaat	atccatcg	tcggccatct	ccaggcggcc	atccggggca	atccgggtca	9360
gctgtgggtt	ctggccaccc	gttcgcgtat	tgctgtcttc	gtcgttgcgt	acccggctgt	9420
gctggccgggg	tttgcattact	ggtagcaga	atgtacatcc	gatacgcg	cgaaagggtaa	9480
gcaactgtc	ctggaaaacg	tgcgcaccc	gagccaaacaa	ataatgttc	ttccgtttcc	9540
gtgtttcgt	aagtctgtt	acgcggaaat	cagegccttgc	caccattat	ttccggatcc	9600
gcatcgccgg	atgtgtgtt	cttccatgtg	gaaacatcc	atctgtat	acgaaggcgt	9660
ggcatgttgc	cttgcgtttt	ttttcttgc	ccggccgtt	ccatccgc	agttgtttatc	9720
ctccaaacg	ttcccaatc	cgggcatgtt	catcatgt	aaacccgtatc	gtgacatcc	9780
tctctgtt	catcggtt	atttccccc	tgaacagaaa	ttcccccttca	cacggggacaa	9840
tcaagtgtacc	aaacaggaaaa	aaaccgcct	taatcgatcc	cgctttatca	gaaggccgac	9900
attaacgccc	ctggggaaaat	taacacgt	ggacgcggat	gaacaggcag	acatotgtga	9960
atcgcttcc	gaccacgtcc	atgcgttta	ccgcacgttc	tcgcgtgttgc	tcgggtatgc	10020
cggtgaaaac	cttcgcaca	tgcatgtttcc	ggagacggtc	acatgttgc	tgtaaaggcga	10080
tgccggggac	agacaagccc	gtcaggggcc	tcacgggggt	gtcgggggt	gtcggggcc	10140
agccatggcc	cagtcaatgc	ggatcgtatgc	agttgtatcc	ggcttacta	tgccgtatca	10200
gagcagatgt	tactgtggat	ggccatata	cggtgtggaa	tacccgcac	atgcgtatgg	10260
agaaaaatcc	gcatcggggc	ctttccgttgc	tccgcgttca	ctgactcg	cgctcggttgc	10320
gttcgggttgc	ggcgagccgt	atcagtcac	tcaaaggcg	taatacggtt	atccacagaa	10380
tcagggttgc	acgcggggaa	gaacatgttgc	gaaaaaggcc	acggaaaaggc	caggaaacgt	10440
aaaaaaggccg	cgttgtgttgc	gtttttccat	aggctccgc	ccccctgcga	gcataccaaaa	10500
aatcgacgtt	caagtccatgc	gttgcgaac	ccgcacaggac	tataaagata	ccaggcggtt	10560
cccccttgcg	gtcccttcgt	gtcgcttc	ttccggaccc	tgccgttcc	cggtatcc	10620
tccgccttcc	tccttcgtgg	aaagctgtgc	tttttcata	gctcagctgt	taggtatcc	10680
agttcggtgt	aggcttcgttgc	cttcacgttgc	ggctgtgtgc	acgaaacccc	cgttcagcc	10740
gaccgcgttgc	cotttacccg	taactatgt	cttgcgttca	acccggtaa	acacgactta	10800
tcgcctactgg	caggccaccc	tggtaacagg	attagcagag	cgaggatgttgc	aggcggtgt	10860
acagatgttc	tgaatgtgttgc	gtcataactac	ggctacatcc	gaaggacatgt	atttgggtatc	10920
tcgcgttgc	tgaaggccat	tacccatcg	aaaagagggttgc	gtatgttgc	atccggccaaa	10980
caaacccatcc	ctggtaggg	ttgttttttt	tttgcgttgc	acgcgtttac	gcccggaaaa	11040
aaaggatcc	tggatgttttgc	tttatccat	tttacgggttgc	ctgacgttca	gtqaaacgaa	11100
aactcacgtt	aagggttttttgc	tttgcgttgc	tttacgggttgc	tttgcgttca	tttgcgttca	11160
ttcttagataa	tacgacttcac	tata				11184

<210> 14
<211> 11184
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid

actcctcgag	ggcgcgcact	agtgggata	cgcgggtat	gcccgtttt	agcatattga	7920
cgaccaatt	ctcatgttg	acagtttac	atcgataagc	ttaatgegg	tagtttattca	7980
cagttaaat	gctaaccgag	tcaggccacg	tgtatgaaat	ctaaatcg	gctcatcg	8040
atccctggc	ccgtcaccc	ggatgtgtt	ggcatagatg	tggatgtcc	ggtaactcg	8100
ggctcttgc	gggatatgt	ccatccccac	agatcgcc	gtcaactatgg	cgtgtgtatg	8160
gctgtatatg	cggtatgtca	attttatgc	gacccgttc	tcggagact	gtccgacgc	8220
tttggcccg	gcccaggatct	gtcgatgtt	ctacttggag	ccactatcga	ctacgcgtac	8280
atggcgacc	caccgcgtt	gttgatgtt	tacggccgac	gcatcggtt	ggcgtatcacc	8340
ggccgcacag	gtgggttgc	ttggcgatct	atccgcgac	tcacccatgg	ggaagatcg	8400
gctgcgcact	tcgggtctat	gagcgctgt	tttgggttgg	gtatgggttgc	aggccccgt	8460
gccggggggac	tggtggggc	catctccgtt	catggccat	tccttgcgg	ggcggtgttc	8520
aacggccctc	acttactatc	gggtgttgc	ctaatgcagg	atgcgtatca	gggagaggt	8580
cgaccgtgc	ccttggagac	cttcaaccac	tcgtatctt	tcgggtggcc	gcggggcatg	8640
actatcgatc	cggcgtttat	gactgttgc	tttatatcgt	aactcgtagg	acagggtccg	8700
gcagcgcttc	gggttatttt	cgccgaggac	cgcttttgtt	ggagcgac	gatgtatcg	8760
ctgtgttgc	cggttatttgc	aaatccgtac	ggccgtgttc	aaccttgcgt	cactgtgtcc	8820
gcccaaacac	gttccggggc	gaaggcggc	attatccggc	gcatggccgc	cgacgcgt	8880
ggctgtatct	tggtgggtt	cgcgacgcg	ggctgtgttgc	cattccat	tatgtattt	8940
otcggttgc	cgggcatcg	gttccggccg	tttgcagggtt	tgtgttccag	cgaggtagat	9000
gacgacatc	aggggacatgc	tcaaggatgc	cttgcgttgc	tttccaggatc	aacttcgtatc	9060
actggacgc	tgatgttcac	ggggatattt	ggccgttccgg	cgaggacatgc	gaacgggtttg	9120
gcatgtatgt	taggcggcc	cttataactt	gtctgttgc	cccggttgc	tcgggtgtca	9180
tggagccgg	ccacccgtac	ctgaatgttgc	gcccggcc	cctcgatca	ggattccatca	9240
ctccaagat	tggacccat	caattdttgc	ggagaatctgt	aatgtcgcc	accacccctt	9300
ggcagacatc	atccatcgcc	tcggccat	ccaggccgc	cacggggcc	atccgggttgc	9360
ggcttgggtc	ctggccacgg	tggtcgatgt	tcgtgttgc	gtcggttgc	acccggctag	9420
gctggccggg	tttgcattatc	gtttagcgtt	atgaatccac	gatagcgcc	cgaaatgttgc	9480
gcaactgtgt	ctgcaaaaac	tctgcgttgc	gagaacaaac	atgaatgttgc	ttcgtgtttcc	9540
gtgttgcgt	aaatgtcgaa	acggcgaaatgt	cgccgcctgc	caccattatgc	ttccggatct	9600
gcatcgatgc	atgtgtgttgc	cttacccgttgc	gaacacatca	atcgatgttgc	acgaaggcgt	9660
ggcatgtacc	ctgtgtgttgc	tttcttgcgtt	cccgccgtat	ccatcagcc	agttgtttac	9720
cctacaacgc	tttccgtatc	cgggcatgttgc	catcatcgat	aaccgttgc	gttgcgtatcc	9780
tcttgcgttgc	catcggttac	attttccatc	tgaacagaaaa	tttccggatca	cacggaggca	9840
tcaatgtacc	aaacacggaaa	aaacccgttgc	ttttatgcgc	cggtttatca	gaaggccatgc	9900
attaacgttc	ctggaaatgc	tcacacgtatc	ggacccgtat	ggacaggatc	acatctgtgc	9960
atcgcttcac	gaccacgtgc	atgagtttac	ccgcgtatgc	ctcgcgttgc	tcgggtatgt	10020
cgtgtaaatc	cttgcgtatca	tgatgttgc	ggagacgttc	acaggttgc	tgtaaatgcgc	10080
tgccggggatc	agacaaacgg	gtcaggggcc	gtcaggccgg	gtcaggccgg	gtcggggccgc	10140
agccatgtacc	cagtgtatgt	cggtatgttgc	atgtgtatgt	ggcttacttgc	tgcggcatca	10200
gaggcatgttgc	tactgtatgt	ggccatcatgc	cggtgttgc	tacccgcacat	atgcgttaatgc	10260
agaaaaatacc	gcatcgatgc	cttccgttgc	tttgcgttca	ctgactcgatc	ggcgttgcgtc	10320
gttgcgttgc	ggcgatgggt	attcgtatcc	tccaaaggccg	taatacggtt	atccacagaaa	10380
tcaggggatc	acggcgaaatgc	gaaatgttgc	gcaaaaggcc	agccaaatggc	caggaaatgttgc	10440
aaaaaaatcc	cggtgttgc	ttttttccat	agggttgcgc	ccccgtatgc	gatcacaatgc	10500
aatcgatgttgc	caagtgtatgt	tggtcgatgttgc	ccggatgtatgc	tataaagatgc	ccaggcgatgc	10560
ccccctggaa	gttccctgttgc	ggcttgcgtt	gttgcgttgc	ttccggatcc	cggtatccgt	10620
tccggcttgc	tttcccttgcgg	aaacgttgcgg	tttttgcata	gctcaccgttgc	taggtatctc	10680
agttcggtgt	agggtcgatgc	cttccaaatgc	ggctgtgtgc	acgaaatccgc	cggttgcgttgc	10740
gaccgttgc	ctttatccgttgc	taaactatgttgc	tttgcgttgc	acccggatgc	acacgtacttgc	10800
tccggatgttgc	caggccgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10860
acagatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10920
tgccgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10980
caaaaaatcc	cttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	11040
aaaggatctc	aaaggatctc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	11100
aacttcgttgc	aaaggatctc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	11160
tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	11184

<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 15
agtatcgtagg tagagagctg c
<210> 16
<211> 36
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer
<400> 16
taatacgact cactataggg atgtggctgg agatgc

21

36

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
1 July 2004 (01.07.2004)

PCT

(10) International Publication Number
WO 2004/055216 A3

- | | | |
|--|---|---|
| (51) International Patent Classification ⁷ : | C12Q 1/70 | (81) Designated States (<i>national</i>): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TI, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW. |
| (21) International Application Number: | PCT/US2003/039722 | |
| (22) International Filing Date: | 12 December 2003 (12.12.2003) | |
| (25) Filing Language: | English | |
| (26) Publication Language: | English | |
| (30) Priority Data: | 60/433,303 13 December 2002 (13.12.2002) US | |
| (71) Applicant (<i>for all designated States except US</i>): | FOX CHASE CANCER CENTER [US/US]; 333 Cottman Avenue, Philadelphia, PA 19111 (US). | |
| (72) Inventors; and | | |
| (75) Inventors/Applicants (<i>for US only</i>): | ZHU, Qing [CN/US]; 1504 Liberty Court, North Wales, PA 19454 (US). GUO, Ju-Tao [CN/US]; 26 Township Line Road, Apt. A7, Elkins Park, PA 19027 (US). SEEGER, Christoph [US/US]; 407 Waring Road, Elkins Park, PA 19027 (US). | |
| (74) Agents: | RIGAUT, Kathleen, D. et al.; Dann, Dorfman, Herrell & Skillman Suite 2400, 1601 Market Street, Philadelphia, PA 19103-2307 (US). | |

Published:

- *with international search report*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report:
5 August 2004

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/055216 A3

(54) Title: REPLICATION OF HEPATITIS C VIRUS IN NON-HEPATIC EPITHELIAL AND MOUSE HEPATIC CELLS

(57) Abstract: Cells and cell lines which replicate HCV of non-hepatic human and non human origin are disclosed. Also provided are methods of using such cells and cell lines to identify anti-HCV agents for the treatment of HCV infection.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/39722

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/70

US CL : 235/5

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 235/5, 235; 435/354, 358, 363

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubMed**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,968,775 A (HOUGHTON et al.) 19 October 1999 (19.10.1999), col. 2, lines 63-67; col. 3, lines 1-5 and 16-28; col. 13, lines 59-67; and col. 14, lines 1-25	1, 3, 7, 9-13, 15
—		—————
A		2, 4-6, 8, 14
X	MACEJAK, D.G. Enhanced antiviral effect in cell culture of type 1 interferon and ribozymes targeting HCV RNA Journ. of Viral Hepatitis June 2001, Vol. 8, pages 401 and 402	16-18

<input type="checkbox"/> Further documents are listed in the continuation of Box C.	<input type="checkbox"/> See patent family annex.
"A" Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"B" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"C" earlier application or patent published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Z" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	
Date of the actual completion of the international search	Date of mailing of the international search report <i>07 JUN 2004</i>
03 May 2004 (03.05.2004)	
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703) 305-3230	Authorized officer James House <i>Janice Ford</i> Telephone No. (703) 308-0196 <i>JF</i>